Journal of Sustainability Science and Management
Volume 17 Number 7, July 2022: 204-225
eISSN: 2672-7226

WATER QUALITY PREDICTION USING LSTM-RNN: A REVIEW

AZHAR JAFFAR[1,2]*, NORASHIKIN M. THAMRIN[2], MEGAT SYAHIRUL AMIN MEGAT ALI[2,3], MOHAMAD FARID MISNAN[4] AND AHMAD IHSAN MOHD YASSIN[2,3]

[1]Jabatan Kejuruteraan Elektrik, Politeknik Ungku Omar, 31400 Ipoh, Perak, Malaysia. [2]College of Engineering, Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia. [3]Microwave Research Institute (MRI), Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia. [4]College of Engineering, Universiti Teknologi MARA (UiTM) Pasir Gudang, 81750 Masai, Johor, Malaysia.

**Corresponding author: azharfar1977@gmail.com*
Submitted final draft: 15 December 2021 Accepted: 9 April 2022
http://doi.org/10.46754/jssm.2022.07.015

Abstract: Water is a critical component of life on this planet, as humans, animals, and plants all rely on water supplies for survival. Regrettably, human activity has contaminated water sources. As a means of remedying this situation, each river should have an intelligent system that continuously monitors the water quality index at each water source. However, testing and establishing a realistic acceptability level for all water quality and control parameters takes time. Some of the data on the parameter's values were also not immediately available. To address this problem, an estimation based on past samples and readings was applied. The process also implemented an exemplary computer system-based method known as an artificial neural network (ANN) Which has evolved into a promising and versatile tool for data prediction over the last few years. This article analyses and compares a host of ANN techniques that have been researched or used in the past to determine water quality. According to this study's review of available literature, there are two types of ANN computation engines: Feedforward network systems and recurrent network systems that can be used for the purposes of determining water quality. Each method has advantages and disadvantages which were studied and this study hypothesises that with right techniques will be able to accurately predict and account for river water quality.

Keywords: Water quality prediction, artificial neural network (ANN), long short-term memory (LSTM), recurrent neural network (RNN).

Introduction

Water shapes the patterns of human settlement, animal migration and vegetation growth and coverage around the world. Access to clean water is crucial for a healthier life. Not only do plants need water to thrive, but plants that live underwater need high-quality water to survive and reproduce.

Unfortunately, human activities have degraded the integrity and cleanliness of water sources. The contaminated water from polluting industries runs directly into the rivers without any treatment before being released into rivers (Kiran Relangi *et al.,* 2019). The use of insecticides and so on in fertilisers dissolves into the soil, which eventually mixes with groundwater and pollutes it. The extra chemicals from the farm are also discharged directly into the drain, which eventually reaches the rivers. Additionally, waste disposal of food and sewage found in sewage from untreated residential areas adds to the pollution of the river from faulty drainage systems.

Without adequate controls and management systems, water pollution will worsen every year. Therefore, research on this subject is of great interest and importance. Each river must have a mechanism to monitor its water quality. According to the recommendations given in the Water Quality Index (WQI), the difficulty of getting acceptable water quality levels encompass several items in the water itself. Among the factors involved are pH levels, turbidity, dissolved oxygen, biochemical oxygen demand, chemical oxygen demand, solids, algae, temperature and others (Cantor, 2019).

However, testing and ascertaining the value of some of these parameters takes time.

It is because some of them must go through a series of chemical laboratory methods. Also, the changing water conditions due to the weather and the environment can have an affect on the readings (Wei *et al.,* 2019). The difficulty in estimating the rate of system change is closely tied to the diversity of parameter elements and feedback from closed-chain systems. All of these contribute non-linear behavioural features to the data produced and make it dynamic and complex to interpret (Pasini, 2015).

The implementation of a robust river water quality prediction system utilising historical samples and reading methodologies as control for this problem has become the main emphasis of this paper. One of the latest technologies in computing systems capable of doing complicated data processing is termed "deep learning." It is one of the branches of a machine learning system and is a technique used in computer systems to substitute human thinking techniques by learning from examples (Najafabadi *et al.,* 2015). This long-standing technology provides the secret to the sophistication of today's automated systems, specifically artificial neural networks (ANN). Over the last decade, several academics have employed ANN to predict and evaluate the water quality in rivers all over the world.

This method can enhance the accuracy and reliability of the predictions made by the ANN system. Using ANN approaches can help save time in decision-making based on the value of measurements made several years ago (Chen *et al.,* 2020).

Artificial Neural Network

The ANN is a model for data processing that mirrors the processing method used by a biological nervous system, such as the brain. In other words, the ANN operates in parallel with the human brain at all times, processing non-linear correlations between inputs and outputs (Seo *et al.,* 2016). Additionally, this technology is capable of processing huge amounts of data, which can only be processed by a computer.

What is an Artificial Neural Network?

The fundamental notion of an ANN is based on a biological brain's neural network and system, which is known to be capable of digesting large amounts of data to develop new information. The human neural network is depicted in Figure 1.

Neurons are made up of billions of nerve cells and are a component of the human brain. According to Figure 1, neurons communicate with hundreds of other cells via axons. Dendrites get signals from a variety of sensory organs. The soma or body of the neuron is related to the dendrites (Grbatinić *et al.,* 2015).

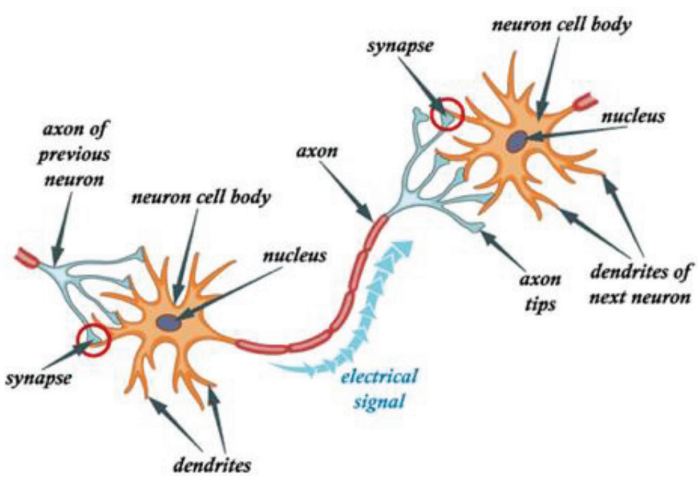

Figure 1: Biological neuron, axons and dendrites in human (Mahanta, 2017)

Dendrites generate electrical impulses that are carried by neural networks. For each of these distinct functions, neurons communicate with another neuron that handles the data. Each axon will form a unique neuron network junction. Synapses connect the axon to other neuron junctions.

Thus, the axon's output signal can penetrate neighbouring neurons. Other neurons will follow suit and the process will be repeated millions of times (M. Islam *et al.,* 2019).

ANN seeks to emulate how humans think and act. There are several main advantages of this, as the ANN system is capable of handling difficult data processing and learning to provide output data that is not just susceptible to the input data provided (Abiodun *et al.,* 2018). Additionally, it does not use data storage-based strategies. Still, it uses the network system itself to prevent data inaccuracy in the event of data loss during processing. With the ability to learn from examples, the ANN system can receive data on current occurrences or real-time events (Abiodun *et al.,* 2018).

Each of these neurons or nodes is a processor that acts on its own. For ANNs to process massive amounts of data, it needs to have a large number of neurons that enable all the processors representing neurons to work in parallel and simultaneously. Figure 2 depicts the mathematical model of a single-layer neural network called a perceptron.

According to Figure 2, x_1, $x2$ till x_m are the network's variable inputs, which come from a variety of sources. Each of the inputs is multiplied with the connection weights, $_{1,2}$ until m or biologically, the synapses. These values indicate the strength of specific nodes or neutrons. A bias function, b is introduced to allow the activation function to change direction. The activation function determines whether to activate the neuron by calculating the sum and bias functions.

It will introduce a nonlinearity value into the output of neurons, as illustrated in Equation 1.

When additional layers of a neural network are used, the system gets more sophisticated. Figure 3 depicts a feedforward multi-layer neural network, also known as a multi-layer perceptron (Vieira *et al.,* 2017) which has four layers in this case.

The first layer is composed of data received from the outside world, i.e., sensor readings. The second and third layers are the hidden or unseen neurons. These neurons are in charge of the

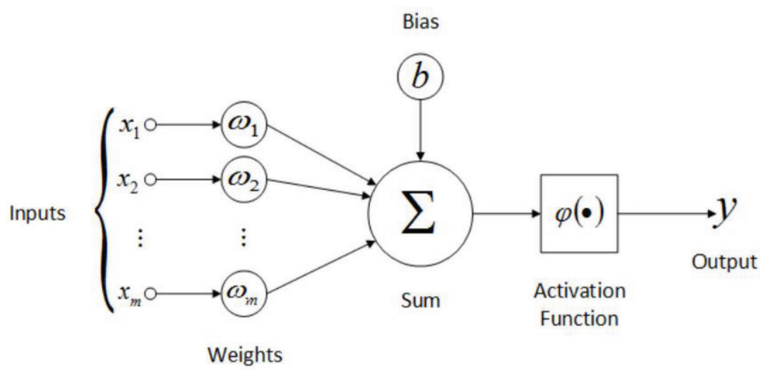

Figure 2: Mathematical model of a single-layer neural network (De Oliveira *et al.,* 2017)

$$b + x_1 \cdot \omega_1 + x_2 \cdot \omega_2 \dots\dots x_m \cdot \omega_m = b + \sum_{n=1}^{m} (x_n \cdot \omega_n)$$

$$y = \varphi(\cdot)(b + \textstyle\sum_{n=1}^{m} (x_n \cdot \omega_n)) \qquad (1)$$

network's internal operations, such as extracting the examined patterns for the system. The final layer is the output layer from the preceding neuro network stage (da Silva *et al.,* 2016).

The more hidden layers in a multi-layer neural network, the more intricate it becomes. However, as the network complexity increases, it delivers more accurate and reliable output.

ANN Basic Architecture

Each artificial neural network architecture has its advantages and disadvantages. The complexity of the design depends on the interconnections between the nodes, their positioning and the composition of the layer.

Feedforward Neural Network

The feedforward architecture is the simplest type of neural network, with only one hidden layer for data processing, as seen in Figure 4. All signal streams flow in the same direction and terminate at the network's output (Hebert *et al.,* 2014). Each perceptron on a layer is connected to a perceptron on the following layer. There are no connections established between perceptrons in the same layer. As a result, the signal flow is forward or "feedforward".

However, due to the large amount of data that must be processed, the majority of feedforward neural network systems feature more than one hidden layer (Bebis, G., & Georgiopoulos, 1994) enabling the simultaneous processing of larger amounts of data. This type of network is referred to as a multi-layer feedforward neural network.

R is the number of inputs, N is the number of hidden layers and S is the number of outputs as shown in Figure 1. The network's input vector

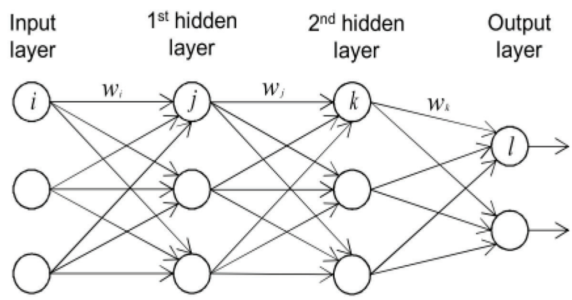

Figure 3: Multi-layer perceptron example (Vieira *et al.,* 2017)

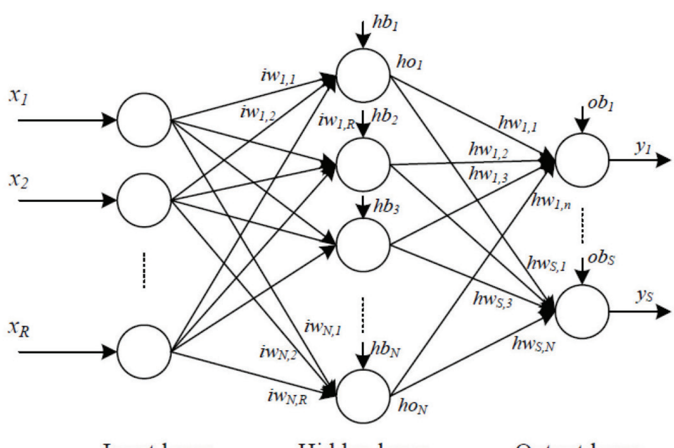

Figure 4: Feedforward neural network (FFNN) with three layers (V. T. Le *et al.,* 2019)

is set to x. In the assessment of water quality, measuring factors such as pH, TE, COD, DO, turbidity and others can be used (Rahmanian *et al.,* 2015). The input weight is *iw* and the hidden weight is *hw*. The bias vectors are denoted by the letters *hb* and *ob*. The output of the hidden layer comes from *ho*. Through *y*, the network output vector is obtained. The input part of the network is represented by the following equations:

$$h_{oi} = f\left(\sum_{j=1}^{R} iw_{i,j} \cdot x_j + hb_i\right), \text{for } j = 1, \ldots, N \quad (2)$$

The output port of the network can be defined using the following equations:

$$y_i = f\left(\sum_{k=1}^{N} hw_{i,k} \cdot x_k + ob_i\right), \text{for } i = 1, \ldots, S \quad (3)$$

Backpropagation is the typical approach for ANN training. This algorithm is used to determine ideal weights by fine-tuning them in response to the preceding iteration's error rate. The network replicates the error derivatives and updates the weights accordingly.

The backpropagation algorithm's purpose is to construct a multilayer feedforward neural network (FFNN) learning algorithm that can be trained to implicitly capture the mapping and obtain a gradient of descent. Training sample data usually makes use of a few simple procedures. The training set is generated using a backpropagation technique. The sample of the input-output iteration is validated using the backpropagation process' performance. This is referred to as the validation set.

When the validation set's performance begins to deteriorate, the iteration process is terminated. The test data is new data that is completely unknown to the network and is used to determine the FFNN model's actual performance (Pasini, 2015). The FFNN processes signals move in a single direction and do not incorporate time dynamics. Multi-level FFNNs can learn more quickly with less output and with fewer weights (Malinowski *et al.,* 1995).

This requires increasing the number of learning cycles and requiring a high level of training accuracy. This makes the network sensitive to weight changes, resulting in convergence issues.

Recurrent Neural Network

Recurrent neural network (RNN)-based systems which are depicted in Figure 5, differ from feedforward because of their temporal dimension. It often reflects on the past and its decision-making depends on what it has learned before.

In an RNN, each output signal from the perceptron is fed back to the hidden layer of the perceptron (Katte, 2018). Recurrent means that the current output will re-enter as an input for future entries. In each successive element, the

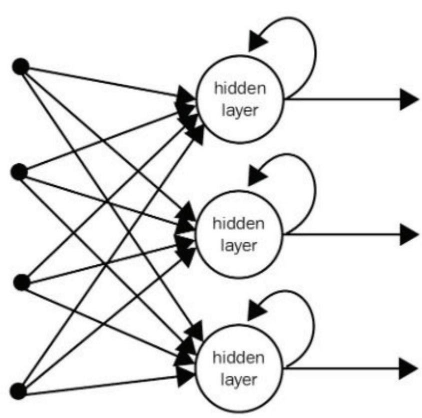

Figure 5: Recurrent neural network (RNN) architecture

system not only takes into account the new input but also remembers the value of the last element (Sherstinsky, 2020). The unfolded recurring input of the RNN is shown in Figure 6. The input is from x and the output is from y through the hidden layer h.

The loop uses W_{hh} and allows information to be passed through one network step to another stage. It is noted that the system has serial entries for each perceptron. The input is connected to a hidden layer by W_{xh}.

From the hidden layer, W_{hy} is the output connection. W_{hh} the weights of the connection are between the hidden layers. X_t represents the current input state while X_{t-1} and X_{t+1}, respectively, represent the input values before and after X_t. Y_t represents the current input state; Y_{t-1} and Y_{t+1} represent the input values before and after Y_t, respectively. RNNs can be used to create multiple output vectors from multiple input vectors.

As with a typical neural network, the output is affected not only by the weight assigned to the input, but also by the "hidden" state vector representing the previous input or output state. As a result, the same input might yield a variety of distinct outputs depending on the previous series' input value.

The input and hidden layer of the network is expressed through the equations below:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1}) \quad (4)$$

To compute the output part of the network can be defined using the following equations:

$$y_t = f\left(W_{hy}h_t\right) \quad (5)$$

The RNN is an ANN model that has been specifically built to deal with long-term dependencies. In actuality, even when the parameters are precisely chosen by humans, RNNs are unable to learn them well. This is referred to as the gradient of vanishing values (Das & Saha, 2019).

The RNN optimises the weights using the gradient technique. Gradients become reduced as the network descends to the lower layer. The gradients will remain constant and will have no information. This modification influences the network's output.

The lack of a difference in the output indicates that the RNN is facing a vanishing gradient, which is not a desirable situation.

The Taxonomy of the ANN

Numerous models of artificial neural networks have been developed to date. Figure 7 depicts an artificial neural network's taxonomy. Numerous data types can be processed using the appropriate network and training methodologies. These networks can be divided into analogue and digital data processing networks. The algorithms for training distinguish between supervised and unsupervised methods. The supervised training

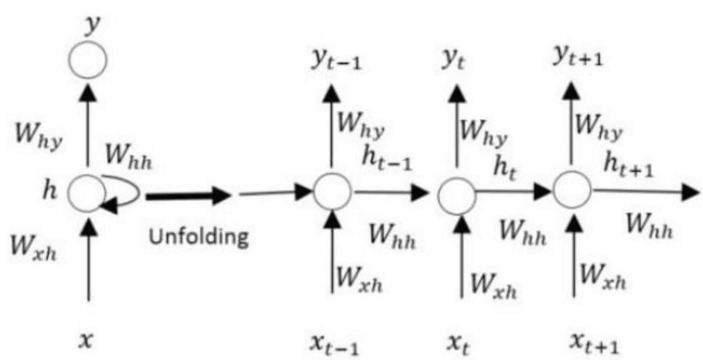

Figure 6: Representation of RNN both in folded and unfolded forms (T. Le *et al.,* 2016)

methods govern the relationship between the ANN model's input and output utilising the training output values.

Unsupervised algorithms attempt to deduce the structure of the data from the data itself. As a result, the supervised method is better suited for function and classification approximation, while the unsupervised method is best suited for grouping tasks.

Feedforward and recurrent neural networks are both examples of supervised learning techniques. As was shown, the feedforward neural network evolves both the Single Hidden Layer Feedforward Neural Network (SLFN) and the Multi-layer Perception Feedforward Neural Network (MLPNN). There are approaches such as Wavelet Neural Networks (WNN) and Extreme Learning Machines (ELM) that are inspired by the SLFN (ELM).

Along with the SLFN and MLPNN techniques, it developed the Radial Function Neural Network (RBFNN) and Probabilistic Neural Network (PNN) methodologies. It also inspired four further techniques, including the Convolutional Neural Network (CNN), the Long Short-Term Memory (LSTM), the Elman Neural Network (ENN) and the Nonlinear Autoregressive Exogenous Neural Network (NARX).

Water Quality Prediction Using ANN

The ANN is advantageous for managing the vast number of inputs required for water quality predictions, which must be flexible in response to changes in the river stream, weather and other variables (Chen *et al.,* 2019). Additionally, it enhances the reliability and evaluation capabilities of the water parameter when compared with the usual approach. It is a significant advantage in terms of resolving the previous model's lack of adaptability and efficiency (Chen *et al.,* 2019). It is hypothesized that applying a non-linear ANN technique will provide an alternative to the forecasting method for water quality.

Khadijah Sulaiman *et al.* (2019) employed ANN to classify the water quality. They employed a basic feedforward neural network to classify the water quality in the Straits of Malacca, Peninsular Malaysia, specifically at the Pontian River, Batu Pahat River and Muar River, based on pattern recognition. Numerous environmental parameters including pH, total suspended solids (TSS), dissolved oxygen (DO),

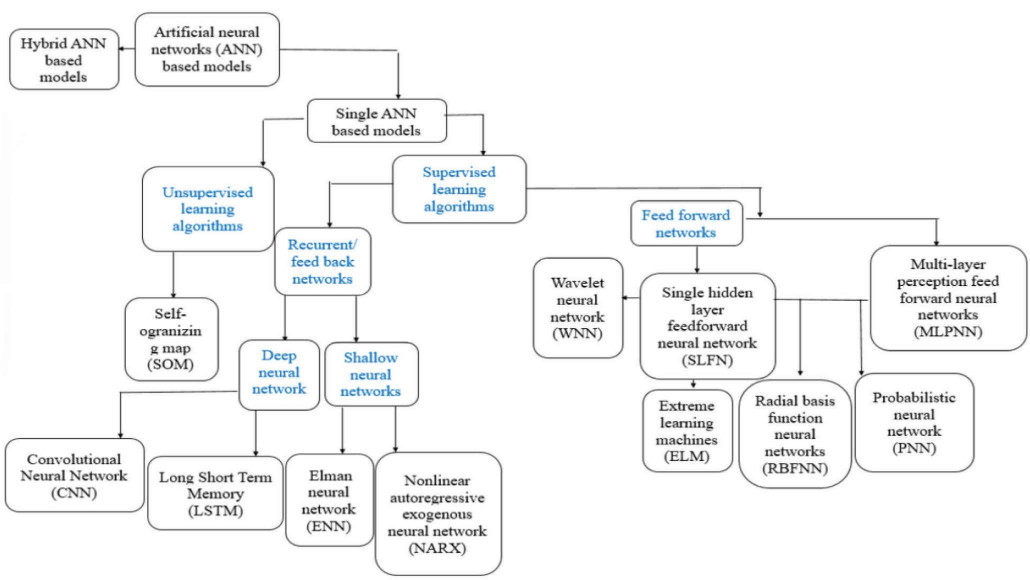

Figure 7: The taxonomy of artificial neural networks (Zhang & Fleyeh, 2019)

chemical oxygen demand (COD), biological oxygen demand (BOD) and ammonia were collected. The study collected 100 data points, of which 80% was training data and 20% were validation data.

Six factors were employed as inputs and five outputs correspond to the five different water quality classifications. The classification accuracy and root mean square error (RMSE) percentages reported in Table 1 are used to calculate the ANN's performance.

20 experiments are given in Table 1 and the ANN accurately assessed the water quality around 16 times, resulting in an accuracy of about 80%. This demonstrates that ANN is a good technique to integrate into a water quality management system.

Archana Sarkara and Prashant Pandey (2015), construct a feedforward neural network as well, but with a more complex multi-layer model based on three-layer networks. The error backpropagation method is included to maximise system efficiency and accuracy (Salami *et al.,* 2016).

The RMSE, correlation coefficient (R) and determination coefficient (DC) were utilised to evaluate the statistical performance of the ANN training in the Yamuna River, India's greatest tributary to the Ganges. The data collection technique entails 72 distinct measurement patterns for the flow discharge, transit length, temperature, pH, electrical conductivity (EC), BOD and DO. 67% of the data sets were utilised for training 33% were used for testing. The findings of the testing are summarised in Table 2.

According to Table 2, ANN-I is a data set that contains data from upstream, central and downstream stations on the Yamuna River. On the other hand, ANN-II includes data from both upstream and downstream stations, whereas ANN-III only includes data from upstream stations. The ANN-II performs the best in terms of training and testing values.

The ANN-I model has an excessive number of inputs, which adds complexity to the model and may result in overfitting the training data (Sarkar & Pandey, 2015). ANN-III requires fewer inputs, which results in poorer overall performance because of a lack of information to define how physical processes work. Simultaneously, ANN-II achieves optimal performance as a result of sufficient input that explains the underlying physical process in detail.

Senlin Zhu and Salim Heddam compared ELM and MLPNN models at the Three Gorges Reservoir in China (Zhu & Heddam, 2020). A forecast of the DO value is compared between various ANN models. In comparison to MLPNN, ELM models with one or more layers of hidden nodes do not require tweaking because of their random generation.

The weight used to connect the input to the concealed node is the unlearned free parameter (Albadr & Tiun, 2017). These concealed nodes

Table 1: Overall results of the testing using ANN

Number of Records	Number of Correctly Classified	Accuracy Percentage
20	16	80.0%

Table 2: Comparative performance of three ANN models

ANN Model	Training (Calibration)			Testing (Validation)		
	RMSE	R	DC	RMSE	R	DC
ANN-I	2.89	.879	.726	3.05	.794	.519
ANN-II	1.71	.907	.822	1.52	.928	.856
ANN-III	2.35	.852	.722	6.91	.654	.283

may be assigned randomly and never updated or they may be inherited in their entirety from their forefathers and mothers. A hidden node's output weight is often learned in a single step, which is comparable to the sum of the steps required to learn a linear model.

The root RMSE, mean absolute error (MAE), R and Willmott index of agreement were used to measure the models' correctness. The trials took place on the Wubu River, the Yipin River, the Huaxi River and a tributary of the Huaxi River. Daily water parameters such as temperature (TE), pH, permanganate index (PI), ammonia nitrogen (NH3-N), EC, COD, DO, total nitrogen (TN) and total phosphorus (TP) are measured.

Approximately 70% of the data is used for training whereas 30% is used for validation. Based on the connections between water quality measurements and DO, the author integrated and assessed nine possibilities as stated in Table 3.

The Wubu River has the most accuracy when the MLPNN4 is employed while the ELM1 is more accurate. MLPNN6 outperformed ELM6 in terms of accuracy for the Yipin River. When ELM2 and MLPNN2 are utilised, Huaxi River achieves the highest accuracy. MLPNN8 is more accurate than ELM6 for the tributary of the Huaxi River. In some combinations, MLPNN outperforms ELM and vice versa.

Although the models used identical input variables for the four rivers, the noteworthy difference is the magnitude of the effect of each independent water quality variable on DO concentrations.

The test findings reveal that when ELM and MLPNN models are used, the system performs well in some rivers but not so well in others because it takes anthropogenic influences into account, the ELM and MLPNN are suitable for estimating DO in low-impact rivers and for struggling in heavily contaminated rivers.

In Vietnam, Thang *et al.* (2019) published research on spatial and temporal monitoring of water quality using RBFNN. Other techniques include decision trees (DT), multilayer perceptron (MLP) networks, Naive Bayes and support vector machines (SVM).

The test was done in the Vietnamese province of Thuan at the Song Quao-Ca Giang (SQ-CG) water system. The RBFNN structure used was a generalisation of the FFNN structure, which consists of a single hidden layer and locally tuned units that are fully connected to an output layer of linear units at a predefined number of degrees (Ahmed, 2017). The hidden layer in RBFNN stimulates the function using a radial basis function derived from the local response. It has several advantageous qualities, including a simple construction, fast training speed and an optimal initial weight dependency (Chen *et al.,* 2019).

Table 3: The input combinations for MLPNN and ELM models

MLPNN	ELM	Inputs Combinations
MLPNN1	ELM1	TE, pH, PI, EC, TP, NH_3-N, TN, COD
MLPNN2	ELM2	TE, EC, TP, NH_3-N, TN
MLPNN3	ELM3	pH, PI, EC, NH_3-N, TN
MLPNN4	ELM4	TE, pH, EC, TP, NH_3-N
MLPNN5	ELM5	EC, TP, NH_3-N, TN
MLPNN6	ELM6	TE, PI, EC, NH_3-N
MLPNN7	ELM7	TE, PI, EC, TP
MLPNN8	ELM8	TE, EC, TP
MLPNN9	ELM9	TE, PI, EC

This paper considered numerous critical parameters, including the pH levels, total suspended solids (TSS), DO, COD, BOD, ammonium (N-NH4+), nitrite (N-NO2-), nitrate (N-NO3-), phosphorous (P-PO43-), TN, TP and Fecal Coliform (FC). Additionally, they are evaluating the water for the presence of heavy metals. A total of 258 samples were analysed from six locations.

The performance of models was assessed using a range of performance metrics, including the percentage of correct and incorrect classifications, the MAE, the RMSE, the relative absolute error (RAE), the root relative squared error (RRSE) and the confusion matrix. The outcome of the research is given in Tables 4 and 5.

As shown in Table 4, RBFNN has the highest accuracy for the spatial test while Table 5 reveals that RBFNN results are somewhat better for the temporal test than the other models, save for the SVM model. The cumulative effect reveals that the RBFNN model outperforms the MLPNN model and may be used to assess water quality.

Probabilistic Neural Networks (PNNs) are a subclass of ANNs that are often used for classification. PPN is a four-layer feedforward neural network architecture. Input, hidden pattern, summation and output are the layers. PNN is a Bayesian classification algorithm that has a high probability of accurately classifying a sample (Chandrasekara *et al.,* 2019).

By making use of PNNs, the necessity for a large dataset during the learning step is eliminated (Yasin *et al.,* 2018). Due to these advantages, training a PNN is rather quick and no learning technique or specified convergence criteria are required (R. Islam *et al.,* 2016). Donya Dezfooli *et al.* (2018) used PNN to classify the Karoon River's water quality. The objectives were to decrease parameter sampling and to allow for rapid classification of water quality.

The author employed nine criteria to determine the quality of the water: BOD, DO, FC, N-NO2, pH, TE, TS, TP and turbidity. As an input and reference, a set of 172 samples and classes from the National Sanitation Foundation Water Quality Index (NSFWQI) is used as input and reference. Two more approaches were chosen for output comparison: The SVM and the K-nearest neighbour (KNN) model.

Around 75% of the datasets were used for calibration and another 25% for testing. The performance of any model is determined by its

Table 4: The percentage of accurate and inaccurate classification for spatial variation

Model	Correctly Classified Samples	Incorrectly Classified Samples
DT(J48)	221 (85.66%)	37 (14.34%)
MLP	217 (84.11%)	41 (15.89%)
Naïve Bayes	206 (79.84%)	52 (20.16%)
RBF	224 (86.82%)	34 (13.18%)
SVM	199 (77.13%)	59 (22.87%)

Table 5: The percentage of accurate and inaccurate classification for temporal variation

Model	Correctly Classified Samples	Incorrectly Classified Samples
MLP	244 (94.57%)	14 (5.43%)
Naïve Bayes	244 (94.57%)	14 (5.43%)
RBF	243 (94.19%)	15 (5.81%)
SVM	198 (76.74%)	60 (23.25%)

error rate (ER), error value (EV) and accuracy as demonstrated in Figure 8.

As illustrated in Figure 8, the PNN model yields the greatest results when calibrating and testing the data. PNN achieved a calibration accuracy of 94.57% and a testing accuracy of 90.70% by utilising only three water quality measures. This demonstrates that PNN has successfully reduced the cost of sampling while increasing the speed of processing for water quality classification.

Nur Suhailayani Suhaimi *et al.* (2020) used Wavelet Neural Networks (WNN) to analyse water quality data from Lake Chini in Pahang, Malaysia. WNN is a combination of mathematical analysis and artificial neural networks (Q. Yang *et al.,* 2017). It is a feedforward neural network with a single hidden layer that uses wavelets as its hidden neurons' activation functions (Harkouss, 2010).

The technique was created for multiresolution signal processing applications such as computer vision, sub band coding, speech and image compression (Teo *et al.,* 2001). Between 2011 and 2015, data collection took place at seven stations in Lake Chini. Researchers measured pH, TE, Optical Dissolved Oxygen (ODO), TDS, turbidity and conductivity. The data set is divided into 70% for training and 30% for testing purposes.

To assess the model's performance, it was compared with a linear artificial neural network and a product-unit neural network (PUNN). Table 6 summarises the investigation's findings.

In Table 6, TT denotes the time in seconds required to complete the sequence. The RMSE is a technique for determining the validity of a model.

True Positives (TP) are classified correctly, whereas False Positives (FP) are classified incorrectly. Due to the complexity of the WNN formula, classifying the quality of the water took longer. Despite this, WNN has a higher accuracy prediction than other models. According to the RMSE analysis, PUNN delivers the best performance model in the quickest amount of time with the least amount of error. By and large, WNN is the best predictor for water quality categorisation due to its increased accuracy.

Takahiro Oga *et al.* (2019) used Convolutional Neural Networks (CNN) to investigate an estimation approach for determining the quality of river water. The estimation is done visually, using the river's colour to determine if it was muddy or clear. The author proposed a new data set and compared it to previously published CNN algorithms such as AlexNet, NIN, GoogleLeNet, VGGNet, ResNet, WideResNet and ResNeXt.

The planned data collection is divided into two categories: "muddy" which refers to blue or green-coloured water and "clear" which refers to brown-coloured water. 400 photographs have been reserved for training purposes while

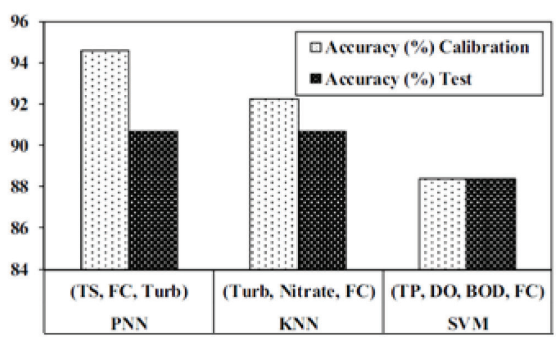

Figure 8: PNN, SVM and KNN models in terms of water quality classification and minimal parameters used (Dezfooli *et al.,* 2018)

Table 6: Prediction results for ANN vs WANN vs PUNN

Data Set (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

50 shots have been reserved for validation purposes. The result is summarised in Table 7.

According to Table 7, "Prop." refers to the proposed data set that includes unique pre-processing data via CNN whereas "Only CNNs" refers to the proposed data set that does not include any further models other than CNN.

All of the results comparing the suggested data to only CNNs and other CNN-based models indicate that the proposed model predicts river water quality more accurately on average.

A Long Short Term Neural Network (LSTM) is a subtype of a recurrent neural network (RNN). What differentiates LSTM from RNN is the presence of feedback connections in LSTM, whereas RNN does not. The LSTM structure enables the model to remember and decide when to forget a state under any time limit (Fan *et al.,* 2020).

Table 7: Test accuracy using data set (%)

	Prop.	Only CNNs	Difference
AlexNet	88.0	71.0	+17.0
NIN	92.0	84.0	+12.0
GoogLeNet	91.0	86.0	+5.0
VGGNet	93.0	85.0	+8.0
ResNet20	91.0	82.0	+9.0
ResNet50	95.0	84.0	+11.0
ResNet101	93.0	70.0	+23.0
WideResNet	92.0	84.0	+8.0
ResNeXt	94.0	84.0	+10.00
Average	**92.1**	**82.2**	**+9.9**

This is done by utilising custom-designed gates and memory cells. The LSTM's long-term memory qualities make it excellent for time series data processing, classification and forecasting (Li *et al.,* 2018). Hu *et al.* (2019) employed the LSTM to forecast pH and water temperature in their smart mariculture project.

Environmental changes such as pH and TE will affect the farmed fish in that experiment. It is vital to have precise data prediction capabilities to expedite the implementation of future countermeasures. The data was collected in Hainan Province, China at the Xincun Town mariculture centre. Pre-processing techniques are used to repair, rectify and de-noise water quality data that has been contaminated by the environment and equipment employed (Apaydin *et al.,* 2020).

Linear interpolation, smoothing and moving average filtering techniques were used to accomplish this. Pearson's correlation coefficient is used to determine the relationship between pH, water temperature and other indicators of water quality.

Using all of this data, an LSTM-based water quality prediction model is created. The LSTM was fed 610 pre-processed samples as input. Table 8 contains the prediction model parameters for TE and pH. The anticipated value is compared to the RNN architecture. As indicated in Table 9, the MAE, RMSE and MAPE were used to evaluate the model's performance.

According to Table 9, the prediction error rate for water quality parameters is less when the RMSE is close to zero (Alsumaiei,

Table 8: Prediction setting for temperature and pH

Item	Temperature Model	pH Model
Input data dimension	5	4
Output data dimension	1	1
Number of hidden layers	15	15
Time step	20	20
Learning rate	0.0005	0.0005
Times of training	10,000	10,000

Table 9: Records of MAE, RMSE and MAPE in long-term prediction

			Training Times	
			500	1,000
MAE	Temperature (°C)	LSTM	0.0421	0.0312
		RNN	0.0439	0.0424
	pH	LSTM	0.0042	0.0325
		RNN	0.0325	0.0052
RMSE	Temperature (°C)	LSTM	0.0519	0.0457
		RNN	0.5340	0.1451
	pH	LSTM	0.6236	0.3108
		RNN	1.0875	0.3254
MAPE	Temperature (°C)	LSTM	0.085	0.052
		RNN	0.078	0.065
	pH	LSTM	0.0092	0.0068
		RNN	0.0102	0.0073

2020). The prediction results for pH and water TE demonstrate that the LSTM model performs better than the RNN model in terms of prediction accuracy and takes less time. The proposed technique has a short-term prediction accuracy of 98.56% for pH and 98.97% for water temperature and long-term prediction accuracy of 95.76% for pH and 96.88% for water temperature.

Another model that is based on RNNs is the Elman Neural Network (ENN). The ENN model preserves the properties of a conventional RNN model but incorporates inputs from the hidden layer to generate a new layer called the context layer (Xu & Zhang, 2019).

In this manner, the context layer acts as a link between the hidden and input levels in terms of feedback. Four levels have been added to ENN: An input layer, a hidden layer, an output layer and a context layer. The context layer is unusual in that it is capable of retaining data from past iterations and using it as input for the current iteration (Network, 2019). In a comparison with other models, ENN is better suited to forecasting time series data. Liu *et al.* (2012) used the ENN model to forecast the DO concentration of Hyriopsis cumingii ponds.

Solar radiation (SR), temperature, wind speed (WS), pH and DO were all used in their investigation. The study collected 816 samples from Singapore's coastal waterways. All the data was standardised prior to training to increase training speed and accuracy. The RMSE, the MAE, the Nash-Sutcliffe coefficient of efficiency and R were used to evaluate the data performance. The findings of this investigation are summarised in Table 10.

In Table 10, a value of zero indicates that the observed mean is an excellent predictor and a good model; a value of one means that the predictor and model are a perfect fit and a value of -1 means that the predictor outperforms the model. Despite the limited sample size, the ENN reveals a moderate correlation between measured and anticipated values. The ANN model has tremendous potential as a forecasting tool and its prediction capability has been demonstrated to be faster than that of a process-based model with few inputs.

The RNN family also includes the Nonlinear Autoregressive Exogenous Neural Network (NARX) architecture. Kazemi *et al.* (2018) released a paper in which they used this model to estimate turbidity in a water distribution trunk in the United Kingdom. Time series analysis and modelling statistical technique that combines an ANN with an autoregressive model with an exogenous (ARX) input (Alsumaiei, 2020).

This enables the capture of non-linear behaviour in an autoregressive time series. NARX is a special case of an RNN in that it models processes utilising lagged input-output variables (Chang *et al.,* 2015). When an exogenous input is added to a NARX network, the network becomes easier to operate because the number of parameters that must be calibrated is reduced. The goal of this study is to forecast a future turbidity event by utilising turbidity and event flow data from the past. The flow is the experiment's input while the turbidity is the experiment's result.

The algorithm employs a single hidden layer of size ten, with input and feedback delays set to three. The performances are evaluated using MAE, NMSE and R. To assess the current case study's performance, the output of NARX is compared to that of the FFN model. Figure 9 illustrates the outcome of the exam. As can be

Table 10: The performance for a different number of the hidden layer

	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
RMSE	0.24	0.29	0.29	0.23	0.29	0.30	0.30	0.30	0.30
MSE	0.23	0.23	0.23	0.29	0.24	0.24	0.24	0.24	0.24
R	0.23	0.20	0.18	0.17	0.16	0.15	0.14	0.14	0.13

seen, NARX provides significantly more insight than the FFN model. For FFN, the normalised mean square error (NMSE) of the trained data (prior events) is 0.559 while for NARX, the NMSE is 0.058.

For the future (predicted) event, MAE was projected to be 0.04 for the FNN model and 0.016 for the NARX model. For the NARX model, FFN has an R value of between 0.66 and 0.97. The NARX network significantly outperforms the FFN network. The trained and projected models exhibit an excellent fit for the NARX-identified occurrences.

The performance of each model is compared in Table 11 using the comparison table. Table 11 discusses ten distinct types of ANN models. Each model has a distinct approach to resolving the water quality issue. A specific model must use pre-processed data to improve the quality of the data entered.

Based on research conducted by researchers that used a variety of methodologies to achieve the maximum level of accuracy possible when forecasting water quality, Table 11 demonstrates the usefulness of the ANN approach in predicting water quality. Numerous comparisons with alternative procedures have also been conducted to demonstrate that the technique used is superior.

The study's overall results, shown in Table 11, indicate that the FFNN technique produces the lowest accuracy of water quality prediction predictions, at about 80%. This is to be expected given that FFNN is the simplest type of neural network, consisting of single-layer nodes that move in a single direction. As a result, critical information from surrounding areas is lost.

Furthermore, the RBFNN approach achieves a spatial accuracy of 86% and a temporal accuracy of 94.19%. Although the RBFNN technique is faster for training, it becomes slower for classification since each node in the hidden layer must analyse each RBFNN vector input (Alexandridis & Zapranis, 2013). This increases the effectiveness of MLPNN over RBFNN, as MLPNN operates globally, with the network output determined by all neurons (Sakaa *et al.,* 2020). Moreover, MLPNN is more capable of identifying fluctuations in data compared with RBFNN (Memarian & Balasundram, 2012).

Additionally, WNN demonstrated somewhat higher accuracy when compared to FFNN. WNN's accuracy percentage ranges between 84.9% and 89.4%. Indeed, WNN is a generalisation of RBFFN and, hence, may be used in place of FFNN (Alexandridis & Zapranis, 2013). Additionally, the PNN approach is based on RBFNN. Donya Dezfooli's (2018) study revealed an optimistic rate of 94.57% on calibration and 90.7% on testing.

When compared with MLPNN approaches, PNN has the advantage of being faster, more accurate and less sensitive to outliers.

In comparison, Takahiro Oga *et al.* (2019) used neural networks based on image processing. The results indicate a comparatively high rate of 92.1%. However, according to the study, it is limited to the image of water turbidity. As a

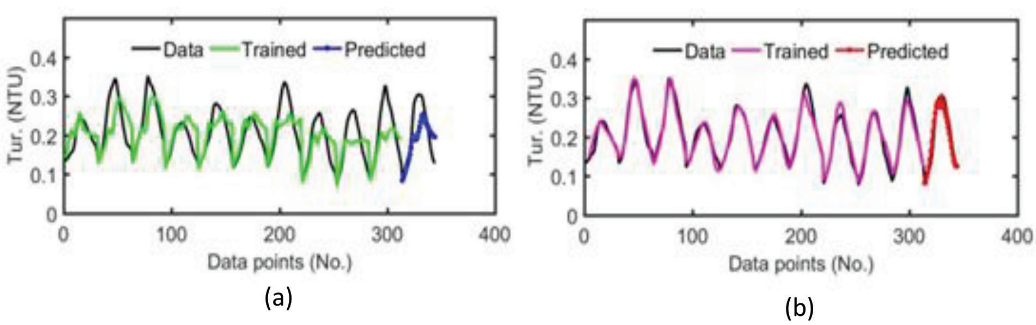

Figure 9: Result of an FFN (a) and NARX network (b) analysis (Chang *et al.,* 2015)

Table 11: Comparison table between various ANN modellings

No.	Model	Input	Data Set	Performance Comparison	Classification Accuracy	Model Accuracy	Previous Work
1.	FFNN	pH, TSS, DO, COD, BOD, ammonia	100 samples 80% - training 20% - testing		RMSE, DC, R	80%	(Khadijah Sulaiman *et al.*, 2019)
2.	MLPNN	EC, DO, BOD	72 samples 67% - training 33% - testing		RMSE, DC, R		(Archana Sarkara & Prashant Pandey, 2015)
3.	ELM	TE, pH, PI, NH3-N, EC, COD, DO, TN, TP	70% - training 30% - testing	MLPNN	RMSE, MAE, R, Willmott index of agreement		(Senlin Zhu & Salim Heddam, 2020)
4.	RBFNN	pH, TSS, DO, COD, BOD, N-NH4+, N-NO2-, N-NO3, P-PO43-, TN, TP, FC	258 samples	DT, MLPNN, Naïve Bayes, SVM	RMSE, MAE, RAE, RRSE, confusion matrix	86.62% spatial & 94.19% temporal	(Viet Thang *et al.*, 2019)
5.	PNN	BOD, DO, FC, N-NO2-, pH, TE, TS, TP, turbidity	172 sample 75% - calibration 25% - testing	SVM, KNN	ER, EV, accuracy	94.57% calibration & 90.70% testing	(Donya Dezfooli *et al.*, 2018)
6.	WNN	pH, TE, ODO, TDS, turbidity, conductivity	70% - training 30% - testing	Linear ANN, PUNN	RMSE, TT, accuracy	84.9% to 89.4%	(Nur Suhailayani Suhaimi *et al.*, 2019)

No.	Model	Input	Data Set	Performance Comparison	Classification Accuracy	Model Accuracy	Previous Work
7.	CNN	Turbidity images	89% - training 11% - validation	AlexNet, NIN, GoogLeNet, VGGNet, ResNet WideResNet, ResNeXt		92.1% avg.	(Takahiro Oga *et al.*, 2019)
8.	LSTM	pH, TE	610 samples		RMSE, MAE, MAPE,	Short term pH-98.56% TE-98.97% Long term pH-95.76% TE-96.88%	(Zhuhua Hu *et al.*, 2019)
9.	ENN	Solar radiation, TE, wind speed, pH, DO	816 samples		RMSE, MAE, R, Nash-Sutcliffe coefficient of efficiency		(Shuangyin Liu *et al.*, 2012)
10.	NARX	Turbidity	1 hidden layer with size of 10	FFN	NMSE, MAE, R		(Ehsan Kazemi *et al.*, 2018)

result, this approach might be deemed relatively ineffective and only suitable for usage in limited circumstances. Based on the results of the last ten trials, the best technique is the LSTM technique.

LSTM investigations typically collect data for both the short and long term. According to Zhuhua Hu *et al.,* the 2019 study offers an extremely positive percentage for both the short and long term.

However, new techniques such as bidirectional LSTM (biLSTM) (Siami-Namini *et al.,* 2019) and gated recurrent unit (GRU) are projected to be able to compete with LSTM techniques due to biLSTM's higher accuracy and GRU's simpler structure (S. Yang *et al.,* 2020).

Conclusion

Water is an incalculably valuable resource and a necessary component of human and ecological life. Each individual is responsible for ensuring the purity of their drinking water. Maintaining and protecting water quality presents various challenges, particularly as the ecosystem continues to grow and develop.

Managing pollution levels via water quality predictions is one of the most effective ways to accelerate the discovery of problems. This is a substantial challenge because it involves a non-linear quantity that is subject to climate and environmental change. An Artificial Neural Network (ANN) is a computer system that emulates the way the brain analysis data to develop algorithms for modelling complex patterns and foreseeing challenges. Numerous ANN architectures have been built to date to address a range of difficult problems.

Based on earlier research, it can be concluded that all ANN architectures can predict water quality, albeit to varying degrees of efficiency, performance and the time taken. The RNN-based LSTM model performs the best, with 96% to 98% measurement accuracy.

Acknowledgements

Thanks to the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA for the technical assistance and financial support given to this research (600-RMC/ LESTARI SDG-T 5/3(125/2019)).

References

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), e00938. https://doi.org/10.1016/j.heliyon.2018.e00938

Ahmed, A. A. M. (2017). Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs). *Journal of King Saud University - Engineering Sciences*, *29*(2), 151-158. https://doi.org/10.1016/j.jksues.2014.05.001

Albadr, M. A. A., & Tiun, S. (2017). Extreme learning machine: A review. *International Journal of Applied Engineering Research*, *12*(14), 4610-4623.

Alexandridis, A. K., & Zapranis, A. D. (2013). Wavelet neural networks: A practical guide. *Neural Networks*, *42*, 1-27. https://doi.org/10.1016/j.neunet.2013.01.008

Alsumaiei, A. A. (2020). A nonlinear autoregressive modeling approach for forecasting groundwater level fluctuation in urban aquifers. *Water (Switzerland)*, *12*(3), 1-16. https://doi.org/10.3390/w12030820

Apaydin, H., Feizi, H., Sattari, M. T., & Colak, M. S. (2020). Comparative analysis of recurrent neural network. *Water (Switzerland)*, *12*, 1-18.

Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. *IEEE Potentials*, 27-31.

Cantor, A. (2019). Water quality parameters. *Water Distribution System Monitoring*, 113-128.

https://doi.org/10.1201/9781439800539.axb

Chandrasekara, V., Tilakaratne, C., & Mammadov, M. (2019). An improved probabilistic neural network model for directional prediction of a stock market index. *Applied Sciences (Switzerland)*, *9*(24), 1-21. https://doi.org/10.3390/app9245334

Chang, F. J., Tsai, Y. H., Chen, P. A., Coynel, A., & Vachaud, G. (2015). Modeling water quality in an urban river using hydrological factors - Data driven approaches. *Journal of Environmental Management*, *151*, 87-96. https://doi.org/10.1016/j.jenvman.2014.12.014

Chen, Y., Fang, X., Yang, L., Liu, Y., Gong, C., & Di, Y. (2019). Artificial neural networks in the prediction and assessment for water quality: A review. *Journal of Physics: Conference Series*, *1237*(4). https://doi.org/10.1088/1742-6596/1237/4/042051

Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences (Switzerland)*, *10*(17). https://doi.org/10.3390/app10175776

da Silva, I. N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., & dos Reis Alves, S. F. (2016). Artificial neural networks: A practical course. *Artificial Neural Networks: A Practical Course*, 1-307. https://doi.org/10.1007/978-3-319-43162-8

Das, T., & Saha, G. (2019). Addressing big data issues using RNN based techniques. *Journal of Information and Optimization Sciences*, *40*(8), 1773-1785. https://doi.org/10.1080/02522667.2019.1703268

De Oliveira, R. M. S., Araújo, R. C. F., Barros, F. J. B., Segundo, A. P., Zampolo, R. F., Fonseca, W., Dmitriev, V., & Brasil, F. S. (2017). A system based on artificial neural networks for automatic classification of hydro-generator stator windings partial discharges. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, *16*(3), 628-645. https://doi.org/10.1590/2179-10742017v16i3854

Dezfooli, D., Hosseini-Moghari, S.-M., Ebrahimi, K., & Araghinejad, S. (2018). Classification of water quality status based on minimum quality parameters: Application of machine learning techniques. *Modeling Earth Systems and Environment*, *4*(1), 311-324. https://doi.org/10.1007/s40808-017-0406-9

Fan, H., Jiang, M., Xu, L., Zhu, H., Cheng, J., & Jiang, J. (2020). Comparison of long short term memory networks and the hydrological model in runoff simulation. *Water (Switzerland)*, *12*(1), 1-15. https://doi.org/10.3390/w12010175

Grbatinić, I., Marić, D. L., & Milošević, N. T. (2015). Neurons from the adult human dentate nucleus: Neural networks in the neuron classification. *Journal of Theoretical Biology*, *370*, 11-20. https://doi.org/10.1016/j.jtbi.2015.01.024

Harkouss. (2010). Accurate wavelet neural network for efficient controlling of an active magnetic bearing system. *Journal of Computer Science*, *6*(12), 1457-1464. https://doi.org/10.3844/jcssp.2010.1452.1459

Hebert, C., Caissie, D., Satish, M. G., & El-Jabi, N. (2014). Modeling of hourly river water temperatures using artificial neural networks. *Water Quality Research Journal of Canada*, *49*(2), 144-162. https://doi.org/10.2166/wqrjc.2014.007

Hu, Z., Zhang, Y., Zhao, Y., Xie, M., Zhong, J., Tu, Z., & Liu, J. (2019). A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture. *Sensors (Switzerland)*, *19*(6). https://doi.org/10.3390/s19061420

Islam, M., Chen, G., & Jin, S. (2019). An overview of neural network. *American Journal of Neural Networks and Applications*, *5*(1), 7. https://doi.org/10.11648/j.ajnna.20190501.12

Islam, R., Musabbir, S. R., Ahmed, I. U., Hadiuzzaman, M., Hasnat, M., & Hossain, S. (2016). Bus service quality prediction and attribute ranking using probabilistic neural network and adaptive neuro fuzzy inference system. *Canadian Journal of Civil Engineering*, *43*(9), 822-829. https://doi.org/10.1139/cjce-2016-0119

Katte, T. (2018). Recurrent neural network and its various architecture types. *International Journal of Research and Scientific Innovation*, *V*(III), 124-129.

Kazemi, E., Mounce, S., Husband, S., & Boxall, J. (2018). Predicting turbidity in water distribution trunk mains using nonlinear autoregressive exogenous artificial neural networks. *3*, 1030-1019. https://doi.org/10.29007/9r3b

Kiran Relangi, N. D. S. S., Chaparala, A., & Sajja, R. (2019). Perspectives in water quality assessment. *International Journal of Recent Technology and Engineering*, *8*(2 Special Issue 4), 7-11. https://doi.org/10.35940/ijrte.B1002.0782S419

Le, T., Kim, J., & Kim, H. (2016). *Analyzing Effective of Activation Functions on Recurrent Network for Analyzing Effective of Activation Functions on Recurrent Neural Networks for Intrusion Detection*. *3*(April 2018), 91-96. https://doi.org/10.9717/JMIS.2016.3.3.91

Le, V. T., Quan, N. H., Loc, H. H., Duyen, N. T. T., Dung, T. D., Nguyen, H. D., & Do, Q. H. (2019). A multidisciplinary approach for evaluating spatial and temporal variations in water quality. *Water (Switzerland)*, *11*(4). https://doi.org/10.3390/w11040853

Li, Z., Peng, F., Niu, B., Li, G., Wu, J., & Miao, Z. (2018). Water Quality Prediction Model combining Sparse Auto-encoder and LSTM Network. *IFAC-PapersOnLine*, *51*(17), 831-836. https://doi.org/10.1016/j.ifacol.2018.08.091

Liu, S., Yan, M., Tai, H., Xu, L., & Li, D. (2012). Prediction of dissolved oxygen content in aquaculture of hyriopsis cumingii using elman neural network. *IFIP Advances in Information and Communication Technology*, *370 AICT*(PART 3), 508-518. https://doi.org/10.1007/978-3-642-27275-2_57

Mahanta, J. (2017, July 10). *Introduction to Neural Networks, Advantages and Applications*. https://towardsdatascience.com/introduction-to-neural-networks-advantages-and-applications-96851bd1a 207

Malinowski, A., Cholewo, T. J., & Zurada, J. M. (1995). Capabilities and limitations of feedforward neural networks with multilevel neurons. *Proceedings - IEEE International Symposium on Circuits and Systems*, *1*(April), 131-134. https://doi.org/10.1109/iscas.1995.521468

Memarian, H., & Balasundram, S. K. (2012). Comparison between multi-layer perceptron and radial basis function networks for sediment load estimation in a tropical watershed. *Journal of Water Resource and Protection*, *04*(10), 870-876. https://doi.org/10.4236/jwarp.2012.410102

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1), 1-21. https://doi.org/10.1186/s40537-014-0007-7

Network, N. (2019). *Runoff Prediction Method Based on Adaptive Elman*.

Oga, T., Umeki, Y., Iwahashi, M., & Matsuda, Y. (2019). River water quality estimation based on convolutional neural network. *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2018 - Proceedings*, *November*, 1305-1308. https://doi.org/10.23919/APSIPA.2018.8659554

Pasini, A. (2015). Artificial neural networks for small dataset analysis. *Journal of Thoracic Disease*, *7*(5), 953-960. https://doi.org/10.3978/j.issn.2072-1439.2015.04.61

Rahmanian, N., Ali, S. H. B., Homayoonfard, M., Ali, N. J., Rehan, M., Sadef, Y., & Nizami, A. S. (2015). Analysis of physiochemical parameters to evaluate the drinking water quality in the state of perak, Malaysia. *Journal of Chemistry*, *2015*(Cd). https://doi.org/10.1155/2015/716125

Sakaa, B., Chaffai, H., & Hani, A. (2020). ANNs approach to identify water demand drivers for Saf-Saf river basin. *Journal of Applied Water Engineering and Research*, *8*(1), 44-54. https://doi.org/10.1080/23249676.2020.1719220

Salami, E. S., Salari, M., Ehteshami, M., Bidokhti, N. T., & Ghadimi, H. (2016). Application of artificial neural networks and mathematical modeling for the prediction of water quality variables (case study: Southwest of Iran). *Desalination and Water Treatment*, *57*(56), 27073-27084. https://doi.org/10.1080/19443994.2016.1167624

Sarkar, A., & Pandey, P. (2015). River water quality modelling using artificial neural network technique. *Aquatic Procedia*, *4*(Icwrcoe), 1070-1077. https://doi.org/10.1016/j.aqpro.2015.02.135

Seo, I. W., Yun, S. H., & Choi, S. Y. (2016). Forecasting water quality parameters by ANN Model using pre-processing technique at the Downstream of Cheongpyeong Dam. *Procedia Engineering*, *154*, 1110-1115. https://doi.org/10.1016/j.proeng.2016.07.519

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, *404*(March), 1-43. https://doi.org/10.1016/j.physd.2019.132306

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The performance of LSTM and BiLSTM in forecasting time series. *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019*, 3285-3292. https://doi.org/10.1109/BigData47090.2019.9005997

Suhailayani Suhaimi, N., Othman, Z., & Ridzwan Yaakub, M. (2020). Analyzing prediction performance between Wavelet Neural Network and Product-Unit Neural Network. *Journal of Physics: Conference Series*, *1432*(1). https://doi.org/10.1088/1742-6596/1432/1/012081

Sulaiman, K., Hakim Ismail, L., Adib Mohammad Razi, M., Shalahuddin Adnan, M., & Ghazali, R. (2019). Water quality classification using an Artificial Neural Network (ANN). *IOP Conference Series: Materials Science and Engineering*, *601*(1). https://doi.org/10.1088/1757-899X/601/1/012005

Teo, K. K., Wang, L., & Lin, Z. (2001). Wavelet multi-layer perceptron neural network for time-series prediction. *Lecture Notes in Computer Science*. http://www.ntu.edu.sg/home/elpwang

Vieira, S., Pinaya, W. H. L., & Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience and Biobehavioral Reviews*, *74*(January), 58-75. https://doi.org/10.1016/j.neubiorev.2017.01.002

Wei, Y., Jiao, Y., An, D., Li, D., Li, W., & Wei, Q. (2019). Review of dissolved oxygen detection technology: From laboratory analysis to online intelligent detection. *Sensors (Switzerland)*, *19*(18). https://doi.org/10.3390/s19183995

Xu, L., & Zhang, Y. (2019). Quality prediction model based on novel elman neural network ensemble. *Complexity*, *2019*. https://doi.org/10.1155/2019/9852134

Yang, Q., Zhang, J., Hou, Z., Lei, X., Tai, W., Chen, W., & Chen, T. (2017). Shallow groundwater quality assessment: Use of the improved Nemerow pollution index, wavelet transform and neural networks. *Journal of Hydroinformatics*, *19*(5), 784-794. https://doi.org/10.2166/hydro.2017.224

Yang, S., Yu, X., & Zhou, Y. (2020). LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp review dataset as an example. *Proceedings - 2020 International Workshop on Electronic Communication and Artificial Intelligence, IWECAI 2020*, *April 2021*, 98-101. https://doi.org/10.1109/IWECAI50956.2020.00027

Yasin, H., Basyiruddin Arifin, A. W., & Warsito, B. (2018). Classification of company performance using weighted probabilistic neural network. *Journal of Physics: Conference Series*, *1025*(1). https://doi.org/10.1088/1742-6596/1025/1/012095

Zhang, F., & Fleyeh, H. (2019). A review of single artificial neural network models for electricity spot price forecasting. *International Conference on the European Energy Market, EEM*, *2019-September*, 1-6. https://doi.org/10.1109/EEM.2019.8916423

Zhu, S., & Heddam, S. (2020). Prediction of dissolved oxygen in urban rivers at the three gorges reservoir, China: Extreme Learning Machines (ELM) versus Artificial Neural Network (ANN). *Water Quality Research Journal*, *55*(1), 106-118. https://doi.org/10.2166/wqrj.2019.053