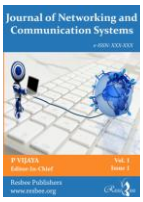
Resbee Publishers

Journal of Networking and Communication Systems

Received 29 April, Revised 14 July, Accepted 23 July



Convolutional Neural Network for Water Quality Prediction in WSN

Suresh Babu Chandanapalli
JNTUK University
Kakinada, Andhra Pradesh, India
sureshbabuchand@gmail.com

Sreenivasa Reddy E
Computer Science & Engineering
Acharya Nagarjuna University
Guntur, Andhra Pradesh, India

Rajya Lakshmi D
Computer Science & Engineering
JNTUK University College of Engineering
Vizianagaram, Andhrapradesh, India

Abstract: By the motivation of applicability of sensor nodes in various appliances like military target tracking, wildlife monitoring, and surveillance, and natural disaster relief, hazardous environment exploration, the incessant monitoring of water quality and features can be too an important technology to scrutinize the physicochemical parameters for increasing the succumbs. For that reason, a diversity of sensors can be positioned in the ponds to gather the need parameters and water quality detection can be performed by exploiting the data classification methods. Here, the Convolutional Neural Network (CNN) is exploited to predict the water quality of in Wireless Sensor Network (WSN). Initially, from the pond, the wireless Sensor Nodes (SN) is exploited to sense the data and CNN is modeled to select the output of the layer. Moreover, exploiting the Firefly with Dual update Process (FF-DUP) is used to choose the routing path in an optimal manner. After that, the outputs of Cluster Head (CH)-based routing protocols are transmitted to the sink node, in that the developed CNN is exploited to classify the water quality parameter. At last, the networking performance of the proposed method is evaluated by exploiting normalized energy utilization with the conventional models such as Low-energy adaptive clustering hierarchy based decision tree (LEACH+DT), LEACH based functional tangent decision tree (LEACH+FDT), fractional artificial bee colony algorithm based decision tree (FABC+DT).

Keywords: Water Prediction; WSN; Cluster Head; Energy; Dead Nodes; Convolutional Neural Network

Nomenclature

Abbreviations	Descriptions
NN	Neural Network
k-NN	K-Nearest Neighbors
SN	Sensor Nodes
DLN	Deep Learning Network
DDM	Data-driven modeling
AI	Artificial Intelligence
CH	Cluster Head
CNN	Convolutional Neural Network
SS	Suspended Solid
SWE	Snow Water Equivalent
AN	Ammoniacal Nitrogen
MFL	Mamdani Fuzzy Logic
BS	Base Station
SGD	Stochastic Gradient Descent

1. Introduction

Water resources pollution and water scarcity have to turn into main issues in the world. To resolve the issue of water resource disparity, in numerous regions, water diversion projects have been modeled [13] [14]. For a water diversion project, water quality is the way into achievement [18]. In the future, at a specific time, the prediction of water quality is to foresee the disparity drift of water environment quality [1]. Moreover, the water quality prediction is of enormous consequence to the control and planning of water quality. To create the arrangement for water pollution control and prevention [16] [17], it is essential to predict the changes in water quality in the future at different pollution levels so that invent a practical plan. For a water diversion scheme, it is further significant to predict the water quality since reasonably an important quantity of the water is transported to solve day by day drinking issues [15]. Hence, in the current society, it is of enormous consequence to discover the approaches of water quality prediction [2].

There is a diversity of algorithms exploited for water quality prediction at abroad and home. [12] These algorithms are chiefly classified into the following four groups such as gray theory, mathematical statistics, and NN and chaos theory. The technique for mathematical statistics is effectual on designing; however, the prediction is not perfect [6]. The algorithm of a gray theory is presently appropriate for approximating exponential functions, not for the complex nonlinear functions. The technique of chaos theory presently can be helpful while the training data is prosperous. The conventional NN and its benefits are an abstraction, non-linearity, self-organization learning is appropriate in order to deal with nonlinear, arbitrarily data. Since the construction of conventional NN [10], it isn't appropriate for processing the time series data.

These algorithms are a method to calculate the parameters of water quality on account of the field data sets and to map the association among the parameter of water quality along with the spatial and temporal variation. Data-driven models denote an extensive variety of models that suggest a system using the data practiced in the real life of the system. DDM is on the basis of evaluating the data distinguishes the system in the study; especially, a design can be stated based on discovering links among the system state variables (internal, input, and output variables) without unambiguous facts of the physical behavior. DDM comprises different classification usually categorized into artificial-intelligent and statistical models that comprise NN, evolutionary computing, fuzzy systems and other areas within machine learning and AI.

The growth and recent development in the amalgamation of diverse AI methods such as genetic algorithm, knowledge-based system, fuzzy inference system and artificial neural network, in water quality modeling, sediment transportation and Dissolved Oxygen concentration was examined by numerous researchers [7].

The main contribution of this paper is to propose CNN for distributed data classification. Using the FFDUP, the optimal cluster head is selected and this is used to minimize the delay and transmission loss. Moreover, the application of CNN is used to classify the aqua status prediction in WSN.

2. Literature Review

In 2018, Segun O. Olatinwo and Trudi-H. Joubert [1], performed a comprehensive study on a few main algorithms to address energy problems in WSN, and enthusiastic to the observed quality of water applications in a technique which contextualizes the considered solution algorithms. Moreover, the solution algorithms were classified, when their weaknesses and strengths were also recognized. Additionally, numerous proposals were done on solution algorithms towards additional enhancement. Hence, this work recommends a practical direction for prospect studies on increasing energy-effectual.

In 2018, Zeshi Zheng et al [2], proposed a method to calculate SWE by the intermission of spatially envoy point measurements by exploiting historical spatial SWE data and k-NN method. It precisely copies measured SWE, by exploiting different data sources for evaluation and training. A k-NN method was used to exclaim data from continuous snow-depth measurements in 10 sensor clusters. Precise SWE estimation against the melt season exhibits the possible to offer the daily, near real-time distributed snowmelt estimates.

In 2019, Ali Najah Ahmed et al [3], addressed an issue in data attained from monitoring stations and experimentation, which were probably polluted using noise signals consequently of systematic and arbitrary errors. Because of the attendance of noise in the data, it was comparatively tricky to create a precise prediction. Therefore, A Neuro-Fuzzy Inference System (WDT-ANFIS) on the basis of the augmented wavelet de-noising technique was recommended and it depended on historical data of the parameter of water quality. The water quality parameters chiefly comprise SS, AN, and pH in the domain of interests.

In 2018, Rodelyn Avila et al [4], worked on the revelation of polluted water whereas boating or swimming or contributing to other recreational activities can reason respiratory and gastrointestinal disease. For water bodies, it was not rare to practice fast fluctuations in water quality, and it was consequently very important to be capable to forecast them precisely and in time in order to minimize the population's revelation to pathogenic organisms.

In 2018, Sadat Mazhar et al [5], worked on pulp and paper industrial wastewater. It was treated with three biological treatments through aerobic, sequential, and anaerobic that was 20 days of anaerobic goes after by 20 days of the aerobic cycle, related with imitation modeling by the MFL model of a few chosen parameters. In sealed plastic bottles, electric air diffusers and least salt medium at control temperature were exploited for anaerobic and aerobic treatments, correspondingly.

3. WSN System Model

Fig. 1 demonstrates the block diagram of water quality prediction utilizing the developed Neural Network in WSN. From Fig 1, the prediction of water quality proposed approach can be described as below:

The numbers of SN are retained within the specific water source. From this specific water source, several parameters of water quality gathered by diverse SN. For instance, the pond one comprises fourteen SN that can gather fourteen types for parameters of water quality from the pond one. In the meantime, from the other water sources, the other SN also gathers the 14 types of parameters of water quality. Moreover, the 5 ponds are exhibited here, in that the 14 types of water quality parameters gathered by 14 sensors. Subsequent to the parameter collection, the SN transmits the information to the sink node by the CH. Subsequently, B is contemplated as BS that is exploited to gather the information from all CH. Initially, the SN are fixed with the water source to extract the parameters of water quality like dissolved oxygen(DO), temperature (temp), turbidity (Turb), pH, electrical conductivity (EC), calcium (Ca2+), nitrate (NO3-), total hardness (TH), ammonia (NH3), total alkalinity (TA), hydrogen sulphide (H2S), carbon dioxide (Co2), biochemical oxygen demand (BoD), diphosphorus (P2). Subsequently, the parameters of water quality are extracted the SN is exploited to transmit the information to BS.

Mostly, when transmitting the data from one SN to another SN, energy utilization is considered as the main issue based on both the amount of transmitting distance and data. To surmount this issue, the CH routing protocol is exploited to transmit the data packets from SN to BS, in that the transfer of data can happen with minimum path exploiting enhanced routing protocol. Here, different CH routing protocols are exploited to transmit the data from SN to BS to surmount confronts on the basis of the energy, delay, power utilization. The CH is exploited to gather all the sensed data from the surrounding SN. Subsequent to gathering data, the data classification can be done on the basis of the neural network classifier in all probable CH. Later than the NN classification [6], the classified data from the different CH is transmitted to BS; in that the classification is done on the basis of the CNN classifier is used to enhance the performance of the classification.

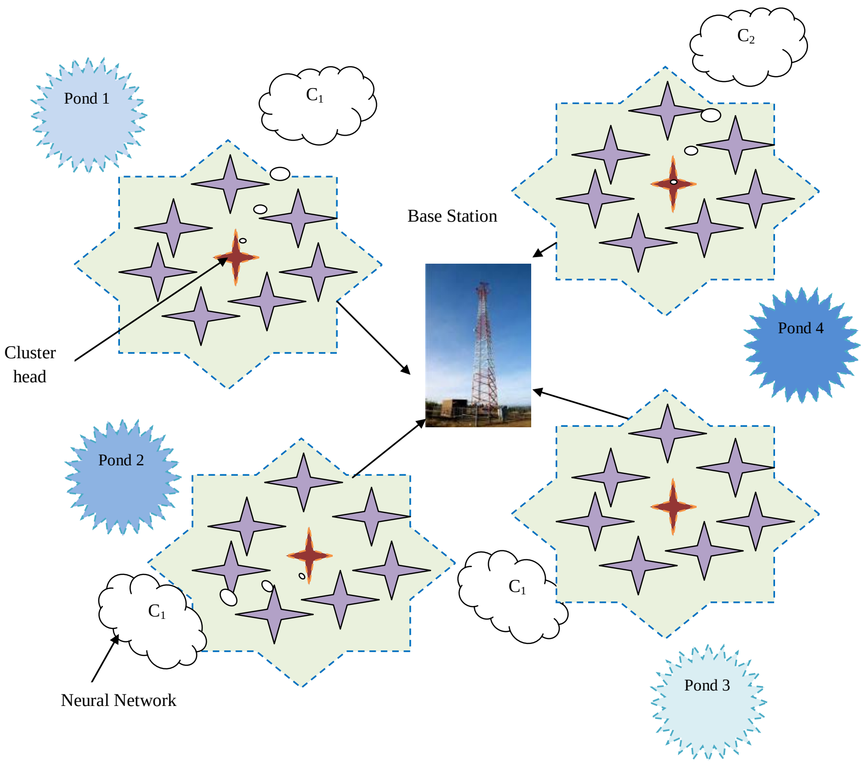


Fig. 1. *System model for water prediction*

3.1 Network Model

The network model [7] of the developed model is shown in this section. This model comprises M a number of SN with only one BS and it is represented as B_s. Here, the M numbers of SN are positioned within the water source to gather the information. Fundamentally, the communications among the SN are obtainable within the radio range on the basis of the wireless link demonstration. Within the radio range to achieve the utmost communication, each and every SN is consistently distributed in the dimensions of L and M meters. Subsequently, the SNs are gathering the water quality from the adjacent pond. Later than gathering the parameters of water quality, the SN is collected into clusters on the basis of its own unique ID of SN. After that, the BS of sink node is positioned in the dimensions of $0.5L,0.5M$, that is fixed closer to the optimal CH position to obtain all the data information from the SNs. Followed by, the CH-based routing method is exploited to transmit the data from each SNs to BS. Mainly, the chosen of the CH is on the basis of the subsequent circumstances: at first, ensure whether the distance between the clusters member with the exacting node is least or not. If the distance is least, so the exacting node is chosen as the CH. Moreover, the CH energy can be computed, in that the computed energy must be as great as probable for CH. In addition, the CH delay must be least to transfer the data. On the basis of the aforesaid 3 constraints, the CH can be estimated to do the superior data transmission. Moreover, from the pond 1, Cluster Head1 is a CH symbol of SN which is exploited to gather the water quality parameters. For that reason, Cluster Head2, Cluster Head3, Cluster Head4, and Cluster Head5 are the CH-based routing protocol for the WSN which comprised the information of the water quality on the basis of pond 2, 3, and 4.

On one occasion, the information is gathered using the CH which is exploited to transmit the information to BS. Moreover, on the basis of the multi-path and free space fading method [8], the data packets are transmitting from normal SN to the cluster node with the energy loss. The transmission of data is done on the basis of the distance among the receiver and the transmitter. After that, the transmitter is exploited to energy dissipation on the basis of the power amplifier and radio electronics. On the basis of the transmission of data distance and nature of the header node and a normal node, the energy is dissipated from each data packet. The energy dissipation of the normal SN can be indicated as below:

$$E_d(N_s) = E_e * D + E_p * D * \left\| N_s^k - CH_j \right\|^4 \tag{1}$$

$$\text{if } \left\| N_s^k - CH_j \right\| \geq d_0 \tag{2}$$

$$E_d(N_s) = E_e * D + E_s * D * \left\| N_s^k - CH_j \right\|^2 \tag{3}$$

$$\text{if } \left\| N_s^k - CH_j \right\| < d_0 \tag{4}$$

In eq. (1), D represents as the data bytes which is transmitted from normal SN, E_p indicates the power amplifier on the basis of the energy parameter, $\left\| N_s^k - CH_j \right\|$ indicates the distance among the normal SN to CH, E_e indicates the electrical energy on the basis of the different factors like filtering, modulation, and amplifiers, that can be indicated in eq. (5).

$$E_e = E_t + E_{da} \tag{5}$$

In eq. (5), E_t represents the energy transmission and E_{da} states the data aggregation energy. While the cluster node CH obtains the D data bytes, the energy dissipation using the CH by receiver stated in eq. (6).

$$E_d\left(CH_s^k\right) = E_e * D \tag{6}$$

Subsequent to transmitting or receiving D data bytes, the energy value in each node can be updated using eq. (7) and (8).

$$E_{t+1}\left(N_s^k\right) = E_t\left(N_s^k\right) - E_d\left(N_s^k\right) \tag{7}$$

$$E_{t+1}\left(CH_s^k\right) = E_t\left(CH_s^k\right) - E_d\left(CH_s^k\right) \tag{8}$$

The aforesaid procedure for the transmission of data is incessant until the dead node development. While the node energy is lesser than 0, that exacting node is denoted as a dead node.

3.2 Routing Protocol

In this section, the complete analysis of the routing protocol on the basis of the FFDUP-enabled CHS is shown. Here, the optimal chosen of CH can be done in 3 manners like initialization, FFDUP for CHS and energy reduction and termination [7].

The steps involved in the initialization procedure are described below:

Initially, in the network area, the SNs are distributed subsequently the energy of the distributed SNs is initialized as E_0.

The centralized clustering on the basis of the routing stratagem is done on the basis of the position of each SN within the network area. Moreover, the position each SN is called at the BS.

Subsequently, the CH configuration can be done using the FF-DUP method that is mostly exploited to choose the best CHs to minimize energy utilization and delay with the best routing.

$$F_{objective} = \beta f^{m} + (1-\beta) f^{n} \,;\, 0 < \beta < 1 \tag{9}$$

$$f^{m} = \alpha_1 * f_{distance} + \alpha_2 * f_{energy} + \alpha_3 * f_{delay} + \alpha_4 * f_{risk} \tag{10}$$

$$f^{n} = \frac{1}{n}\sum_{p=1}^{n}\left\| M_p^{norm} - B_s \right\| \tag{11}$$

The aforesaid procedure is repeated until the terminal obligation. Subsequent to choosing the CH, that is exploited to update the energy of the SN. Once the CH is recognized, the data transmission between the normal SNs and CH is performed. Subsequently, the energy of each node is updated on the basis of each byte of data transmission. The aforesaid 2 processing stages are incessant until each node is represented as dead nodes.

4. Hierarchical Distributed Data Classification Using CNN Model

The developed hierarchical distributed data classification on the basis of the CNN model [10] is exhibited in this section. Here, the distributed data classification can be done in five types at CH, transmit NN to base station and the last classification step in the base station by the amalgamation of different NN.

CNN is a multi-layer network model, which developed from the conventional NN model, which is called as the multi-layer perceptron. CNN mostly comprises the convolution layer, pooling layer, input layer, output layer, and, full connection layer, [11]. It can perform feature extraction and mapping by quick training, and have possessed prediction accurateness; consequently, it is frequently used for the prediction and classification. Here, the structure of conventional NN is enhanced, and a learning network with two hidden layers is modeled.

(i). Data Input Layer

Initially, the preprocessed signal samples are exploited as the input data of the two-layer convolution network, the sample dimension of a single novel signal is 16000*1, and the first feature extraction, is performed by PCA [9]. Moreover, the data samples are predicted to the feature subspace of 784*1, in order that the learning competence and computational complexity of the DLN can be minimized using the dimensionality decrease processing.

(ii) Convolution Layer

The data sample vector of the minimized dimension is transformed into the 28*28 matrix as the input feature graph of the two -layer hidden learning network to perform convolution computation and the local feature extraction is additionally performed. In the learning network, the main convolution layer is collected of 32, 5*5 convolution kernels; the second convolution layer is collected of 64, 5*5 convolution kernels. The formulation of the layer is stated in eq. (12).

$$x_i^m = f\left(\sum_{j=1}^{L_i} w_{j,i} \times y_j^{m-1} 1 + a_i^m\right), i = 1,2,\ldots,N \tag{12}$$

In eq. (12) $w_{j,i}$ indicates the weight value of the convolution kernel for the j^{th} feature graph of the m layer and the i^{th} feature graph of $m-1$ the layer; x_i^m indicates the j^{th} feature graph of the m layer. N indicates the number of the feature graphs of the m layer a_i^m indicates the offset of the i^{th} feature graph of the m layer; $f(\cdot)$ indicates the activation model. In this model, the relu model is stated in eq. (13).

$$f(y) = \max(0, y) \tag{13}$$

iii) Pooling Layer

The pooling layer termed the downsampling layer mostly minimizes the input feature graph dimension and does not alter the count of feature graphs. In this model, the first pool layer exploits the 2*2 filter and the step length is two, as a result, the sample dimension is 14*14 subsequent to the downsampling. The second pool layer exploits the 2*2 filter and the step length is two, as a result, the sample dimension is 7*7 subsequent to the second downsampling. The formulation of the layer is eq. (14)

$$x_i^m = f\left(down\left(x_i^{m-1}\right)\right) \tag{14}$$

Eq. (14), x_i^m, x_i^m indicates the feature graphs of the m and $m-1$ layers correspondingly. $down(\cdot)$ indicates the downsampling function, In this model, it exploits the utmost pooling function. $f(\cdot)$ indicates the activation model, the experiment exploits the Relu model.

iv) Full Connection Layer

The feature graphs sample is combined subsequent to the pooling and convolution, and, the features which are the majority favourable to classification sample are extracted. The one-dimensional feature vector is exploited as the input of the full connection layer, consequently, it is essential to carry out the flatten operation on the two -dimensional output feature graphs of the upper layer. In this model, the number of neurons linked to the two layers is 3164, 1000 correspondingly, equivalent to the dimension of feature graphs of the upper layer. The output of each neuron is stated in eq. (15).

$$h_{w,a}(y) = f\left(w^T y + a\right) \tag{15}$$

In eq. (15), $h_{w,a}(y)$ indicates the output of the neuron, w indicates the weight of the neuron, y indicates the input eigenvector of the neuron and a indicates the offset.

v) Output Layer

From the learning network, the features extracted are input to the softmax classifier to output the decision. Moreover, the output is the probability here the sample is recognized as a class. The signal sample set has a total of 8 classes, the probability of the sample j referred to as the class i is by exploiting the softmax classifier, and the probability in which the sample belongs to each class is computed. At last, the class with the highest equivalent value of the probability is decided as the sample class label.

vi) Network Training

By reducing the loss function the network is trained worldwide, and the SGD model is exploited to recognize the optimal loss function. In the experimentation, the parameters of SGD are set, the learning rate is 0.01, the delay of each learning rate updating is 0.001, and the momentum is 0.9. In the progression of network training, Batch Normalization is exploited to go faster the convergence speed of the training, and the size of each batch is set to 50.

Subsequent to creating the NN in the CH, the data classification can be done. Following executing the classification of data, the classified data is transmitted to the base station from the various CHs. At last, the sink node is exploited to data classification on the basis of the CNN classification.

Next to generating the optimal CH, the NN-based hierarchical distributed classification algorithm is exploited to classify the group of input data. Followed by, the classified outcomes from the number of optimal CH are combined and transmitted to the sink node.

5. Result and Discussions

5.1 Experimental Set up

In this section, the experimentation procedure of WSN with the proposed FFDUP-CNN method was demonstrated. When doing the experimentation procedure, the SN was fixed in the range of 100 m x 100 m. Subsequently, the BS was positioned in the center area of the SN. The network parameters like electrical energy E_e, data aggregation energy E_{da}, transmitter energy E_t and the power amplifier based energy E_p are fixed as like [7] for the simulation of the proposed routing protocol.

5.2 Performance Analysis

In this section, the analysis on the basis of the number of nodes exploiting three parameters like a number of dead nodes, energy consumption, and precision are shown. Moreover, 50 rounds of operation are taken, in that the average utilization can be analyzed concurrently.

Fig 2 shows the analysis of the proposed and existing methods regarding the number of dead nodes. Here, the 21% better than the LEACH+DT, 26% better than the LEACH+FDT and 28% better than the FABC+DT at 50 rounds. Fig 3 exhibits the evaluation of the proposed and existing techniques regarding the number of average energy. Here, the 16% better than the LEACH+DT, 17% better than the LEACH+FDT and 12% better than the FABC+DT at 50 rounds. The performance of the proposed model is superior to the conventional methods such as FABC+DT, LEACH+FDT, and LEACH+DT.

Fig 4 exhibits the analysis of the proposed and existing techniques regarding the number of accuracy. Here, while comparing with the conventional method the proposed algorithm attains the high accuracy as 82% and 83% on the basis of the number of available SNs as 100 and 50 correspondingly.

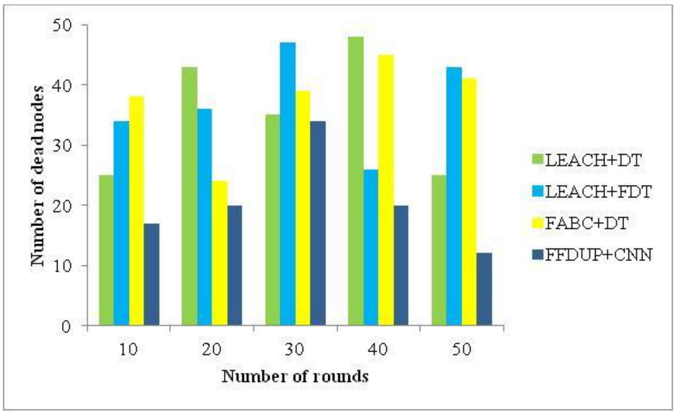


Fig. 2. *Analysis of the proposed and existing method regarding the number of dead nodes*

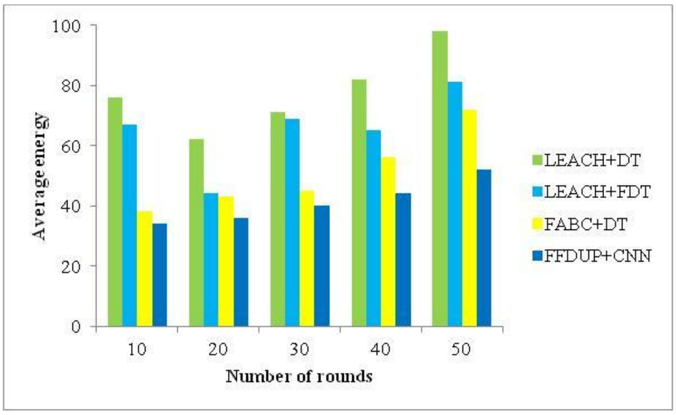


Fig. 3. *Analysis of the proposed and existing technique regarding average energy*

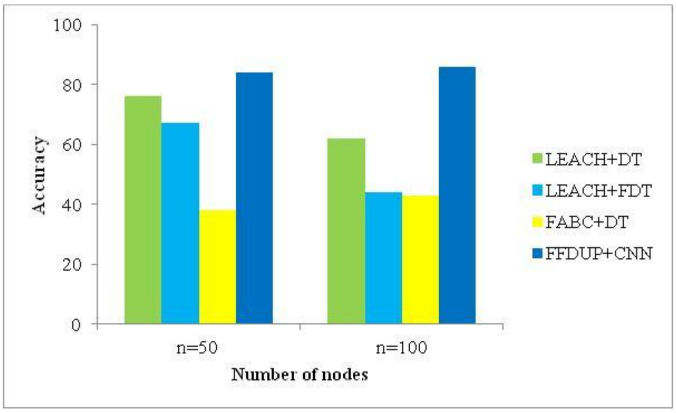


Fig. 4. *Analysis of the proposed and conventional technique regarding the accuracy based on the number of nodes*

6. Conclusion

In any aquatic system analysis, the modelling water quality parameters are of substantial consequence. The conventional modelling technologies are dependent on datasets which include large amount of unknown or unspecified input data and usually comprise of time-consuming processes. In this paper, CNN was used to predict the quality of the water in WSN. Initially, in ponds, the number of sensors is positioned to gather the required water quality parameters, like DO, Ca2+, temp, P2, pH, Turb, EC, TH, NO3−, NH3, H2S, TA, BoD, and Co2. Moreover, fourteen parameters of water quality that can be gathered from 21 ponds were considered by research experts. Subsequent to sensing the parameters of water quality, the neural network is created. As a result, the merging procedure was performed on the basis of the cluster head-based routing protocol.

References

[1] S. O. Olatinwo and T. Joubert, "Energy Efficient Solutions in Wireless Sensor Systems for Water Quality Monitoring: A Review," IEEE Sensors Journal, vol. 19, no. 5, pp. 1596-1625, 1 March1, 2019.

[2] Zeshi Zheng, Noah P. Molotch, Carlos A. Oroza, Martha H. Conklin, Roger C. Bales,"Spatial snow water equivalent estimation for mountainous areas using wireless-sensor networks and remote-sensing products",Remote Sensing of Environment, Volume 215, 15 September 2018, Pages 44-56.

[3] Ali Najah Ahmed, Faridah Binti Othman, Haitham Abdulmohsin Afan, Rusul Khaleel Ibrahim, Ahmed Elshafie,"Machine learning methods for better water quality prediction",Journal of Hydrology, Volume 578, November 2019.

[4] Rodelyn Avila, Beverley Horn, Elaine Moriarty, Roger Hodson, Elena Moltchanova,"Evaluating statistical model performance in water quality prediction", Journal of Environmental Management, Volume 206, 15 January 2018, Pages 910-919.

[5] Sadat Mazhar, Allah Ditta, Laura Bulgariu, Iftikhar Ahmad, Ata Allah Nadiri,"Sequential treatment of paper and pulp industrial wastewater: Prediction of water quality parameters by Mamdani Fuzzy Logic model and phytotoxicity assessment",Chemosphere, Volume 227, July 2019, Pages 256-268.

[6] L. Guo, F. Chen, Z. Dai and Z. Liu, "WSN Cluster Head Selection Algorithm Based on Neural Network," 2010 International Conference on Machine Vision and Human-machine Interface, Kaifeng, 2010, pp. 258-260.

[7] A Shankar, N Jaisankar,"Security enabled cluster head selection for wireless sensor network using improved firefly optimization",International conference on soft computing and pattern recognition, 176-192.

[8] Enrique J. Duarte-Melo, Mingyan Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks", In proceedings of international conference on IEEE Global Telecommunications, vol. 21, pp. 21 - 25, 2002.

[9] W. Puyati and A. Walairacht, "Efficiency Improvement for Unconstrained Face Recognition by Weightening Probability Values of Modular PCA and Wavelet PCA," 2008 10th International Conference on Advanced Communication Technology, Gangwon-Do, 2008, pp. 1449-1453.

[10] Y. Xiao and J. Keung, "Improving Bug Localization with Character-Level Convolutional Neural Network and Recurrent Neural Network," 2018 25th Asia-Pacific Software Engineering Conference (APSEC), Nara, Japan, 2018, pp. 703-704.

[11] W. Shen and W. Wang, "Node Identification in Wireless Network Based on Convolutional Neural Network," 2018 14th International Conference on Computational Intelligence and Security (CIS), Hangzhou, 2018, pp. 238-241.

[12] Nazleeni Samiha Haron, Mohd Khuzaimi B Mahamad, Izzatdin Abdul Aziz, Mazlina Mehat, "A System Architecture for Water Quality Monitoring System Using Wired Sensors", In Proceedings of International Symposium on Information Technology, vol. 4, pp. 1 - 7, 2008.

[13] Li Nan, Wang Ruimei, Zhang Jian, Fu Zetian, Zhang XiaoShuan,"Developing a knowledge-based early warning system for fish disease/health via water quality management", Expert Systems with Applications, vol. 36, pp. 6500–6511, 2009.

[14] Fabrice Rossi, Nathalie Villa, "Support Vector Machine For Functional Data Classification", Neurocomputing, Vol. 69, no. 7–9, pp.730–742, March 2006.

[15] Pawlak, Z.: “Rough Sets and Intelligent Data Analysis”, Information Sciences, vol. 147, pp. 1–12, 2002.

[16] Thangavel, K., Pethalakshmi, “A.: Dimensionality Reduction based on Rough Set Theory: A Review”, Applied Soft Computing, vol. 9, pp. 1–12, 2009.

[17] Sami G. Daraigan, Ahmed S. Wahdain, Ahmed S. Ba-Mosa, Manal H. Obid, "Linear correlation analysis study of drinking water quality data for AlMukalla City, Hadhramout, Yemen", international journal of environmental sciences, vol. 1, no 7, 2011.

[18] W.Brajula and Praveena S,"Energy Efficient Genetic Algorithm Based Clustering Technique for Prolonging the Life Time of Wireless Sensor Network",Journal of Networking and Communication Systems (JNACS),Volume 1, Issue 1, October2018.