Journal of Theoretical and Applied Information Technology
15th October 2022. Vol.100. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A MINING FRAMEWORK FOR REAL BURST LOCATION ESTIMATION AND PORTABILITYOF THE WATER USING DEEP LEARNING

P.VASANTHSENA[1] DR.P.SAMMULAL[2]

[1]Scholar R&D of JNTUH, Asst Professor, Chaitanya Bharathi Institute of Technology, Hyderabad, India
[2]Professor, Department of CSE, JNTU Jagitial, Telangana, India

E-mail: [1]vasanthsena_it@cbit.ac.in/vasanthmtechsit521@gmail.com, [2]sammulalporika@gmail.com

ABSTRACT

Good health policy requires that all people have access to safe drinking water as a basic human right. In terms of national, regional, and local health and development, this is critical. Water and sanitation improvements have been found to provide a positive return on investment in certain areas, since the reductions in health risks and medical expenses much surpass the costs of making the improvements. To check whether that water is safe or not we have some parameters which need to be checked like pH value which ranges from 6.52 – 6.83 and Hardness, Chloramines, Sulphate, Conductivity, Organic carbon, Trihalomethanes, Turbidity, and at last portability. When we acquire a result of 1, we know that the water is safe to drink. If we get a portability value of 0 it is not safe for water consumption Before checking the quality of water, we need to collect all water bodies' images from Google earth maps and mask them and check their pot ability.The project involves data analysis of the different parameters which are involved in checking the portability of water with proper dataset using data processing methods. Random Forest, Decision Tree and other machine learning algorithms are used to make predictions. With the use of VGG image Annotator and leakage location estimate algorithms such as cross correlation of sinusoidal waves and water bodies are masked out of water distribution pipes.

Keywords: *Deep Learning, water portability, image masking, Google earth map, image processing.*

1. INTRODUCTION

The introduction parts contain the overview of the project and the factor for motivation for the project, the main objective of the paper, and the problem statement of the research work.As a fundamental human right, access to clean drinking water is a crucial component of good health policy. In terms of national, regional, and local health and development, this is critical. Water and sanitation improvements have been found to provide a positive return on investment in certain areas, since the reductions in health risks and medical expenses much surpass the costs of making the improvements. The motive is to provide safe drinking water, we need to check the quality of water every 6 months this can be possible only if we have water bodies data. We need to collect all the water bodies of a particular city and check whether they are potable or not. The main objective is identifying water bodies, we need to check the water quality and whether it is safe for drinking or not for that we use a water quality management system. Water quality can be checked by different types of metrics like pH values which range from 6.52-6.83 according to WHO standards and Hardness, Solids, Chloramines, Sulphate, Conductivity, etc. As a last step, we must determine whether or not the liquid is drinkable (a score of 1 indicates it is) or not (a score of 0 indicates it is not). The Problem statement extractsthe nearest water resources of a particular city and identifies the portability of that particular water is useful for drinking or not.

2. RELATED WORK

The suggested system includes a water flow detector. Water detectors detect leaks in the pipeline and notify the controller if they detect water flow. The resistance value shown in the mobile app may be utilised to pinpoint the precise position. The gaps are created by the inability to examine online server data in real time using artificial neural networks [1, 21]. A sensor network design approach is discussed in this paper. The pressure data is gathered using Arduino and XBee (Zigbee), and the exponential curve fitting

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

technique is used to locate and identify the leak. This method is incapable of predicting the size of a leak [2]. In the suggested system, an ultrasonic flowmeter is employed to collect pipe flow data. This system employs NB-IoT and the One-Class-outlier SVM detection technique to assess leakage. The approximate location of the leaking is shown [3]. Leaks are found and located using SVM analysis of pressure and flow data at pipeline connectors and pipes in the planned work. This idea might be enhanced by using cloud services and creating a mobile app [4]. A neural network approach was used to examine pressure changes in pipes in order to detect leaks. [5] To find a leak in our proposed system, we used a weighted average localization technique. This article focuses on the linear pipeline, which excludes alternate pipeline possibilities. The smart water meters were used in smart cities for sample gathering and used for analysis by various deep learning approaches [22].

2.1 Sections and Subsections

Sections and subsections should be numbered and titled as 1.0, 2.0, etc. and 1.1, 1.2, 2.1, 2.2, 2.2.1, etc. Capital letters should be used for the section titles. For subsections, the first letter of each word should be in capital letter and followed by small letters. One line space should be given above the sub section while no space should be given below the heading and text

2.2.2 Identification of sub subsections

Subsub section has to be in sentense case with no spacing above or blow the srat of it.

3. METHODOLOGY

Basically, in the existing system, we need to go to that particular location and take the sample water and test it and check the quality of water whether it is suitable for drinking or not. The major drawback is if we want to check the quality of water at more than 100 locations then we cannot visit each place and take a sample of water to test which is a time-consuming process. Instead of visiting every location we can capture images of water bodies using Google Earth Image and can also check the quality of water using a quality management system. Automated implementation of complex functions is possible. It is possible to reduce the processing time, it's simple to use and maintain. There are less chances of making mistakes. The new system is faster and more accurate than the current one, and it is easier to handle reports.

3.1 Edge masking using Canny Edge Detector Algorithm:

A clever edge detector is used for masking. With the Canny edge detector, a multi-stage algorithm may be used to identify a broad variety of pictures' edges. The image's true edges should be included in the strong edge pixels. Weak edge pixels may be extracted from actual edges or noise/color variations, despite much debate. To get an accurate result, it is necessary to eliminate the weak edges created by the latter explanations. While noise responses are interconnected, the weak edge pixels created by real edges are usually linked to the strong edge pixels. Using blob analysis, a weak edge pixel and its 8-connected neighboring pixels are used to monitor the edge connection. In the event that one of the edges is weak, the weak edge point may be selected as the one that should remain in the blob. In the end, we get G=sqrt(Gx2+Gy2) may be used to approximate a gradient. Angle=inv tan (Gy/Gx) G=gradient approximation for identifying the edge direction. It is possible to alter the Canny algorithm's calculation time and efficacy by adjusting several parameters. Gaussian Filter Size: The first-stage smoothing filter has a direct impact on the Canny algorithm's output. Filters with a smaller effective area produce fewer artefacts and make it easier to see finer details. More blurring occurs when using a bigger filter, spreading out the values of individual pixels more widely throughout the picture. In order to identify bigger, smoother edges (such as the edge of a rainbow), it is better to utilise blurring radii that are greater. A single threshold technique has limitations, but the usage of two thresholds with hysteresis provides for more flexibility. Too high of a criterion might miss out on critical data. It's also possible to overestimate the importance of irrelevant information (such as noise). Defining a universal threshold for all photos is tough. There is currently no proven solution to this issue.

3.2 Potable water Dataset Features with their range:

The pH value of water determines its acid-base equilibrium. The World Health Organization (WHO) permits a maximum Ph level of 8.5 between 6.5 and 8.5. The investigation was restricted to this geographical region. Water hardness is determined by the quantity of calcium and magnesium salts present. One factor influencing raw water hardness before treatment is the amount of time water spends in contact with the hardness-producing substance. When it comes to inorganic or organic salts and minerals, water is

capable of dissolving a wide range of them. These minerals affected the taste and look of the water by diluting its color. When it comes to water usage, this is a significant concern.The greater the TDS value, the more mineralized the water. The highest TDS level for drinking is 1000 mg/liter, which is the optimal limit. Chlorine and chloramine are the most common disinfectants in municipal water systems. Chloramines are formed when ammonia is used to disinfect drinking water. Up to 4 ppm of chlorine in drinking water is deemed safe (milligrams per liter).
Sulfate: Minerals, soil, and rocks all include sulphates, which are naturally occurring compounds. Our air, water, food and soil may contain them. The chemical industry is the primary user of sulphate, which is also found in agriculture. Sulfate concentrations in seawater average 2,700 mg/L. Concentrations typically range from 3 to 30 mg/L in most freshwater, however some reach considerably higher levels (up to 1000 mg/L). Pure water is more of an insulator than a conductor in terms of electrical conductivity. Water's electrical conductivity rises in direct proportion to its ion concentration. Water's electrical conductivity is determined by the concentration of dissolved particles. When it comes to conducting electricity, conductivity (EC) is a measure of a solution's ability to do so.
EC values should not go higher than 400 S/cm, as recommended by the World Health Organization (WHO). NOM and man-made components make up the total organic carbon (TOC) found in source waters. When calculating TOC, the amount of carbon in organic compounds is taken into account. TOCs should not exceed 2 mg/L in treated water and 4 mg/L in source water utilised for treatment, according to the US EPA. Chlorinated water may include trihalomethanes, or THMs, which are compounds that may be present. In drinking water, THM levels vary with the quantity of organic matter present, the amount of chlorine added during treatment, and the temperature of the water being treated. Water with THM concentrations of 80 ppm or less is deemed safe to consume. The quantity of suspended solid matter in a given sample determines water turbidity. Measurement of the water's light-emitting qualities serves as a means of determining the quality of waste discharge. The WHO-recommended turbidity level of 5.00 NTU was not met by Wondo Genet Campus' mean turbidity value of 0.98 NTU.
Portability: Whether or not water is safe for human consumption is indicated by a value of 1 (potable) or 0 (unsafe). Parts per million are used as acronyms in this context. Mg/L is milligrams per liter, whereas g/L is micrograms per liter. g/L water's acidity or alkalinity (0 to 14). Water hardness is measured in milligrams per liter. Particulates: Dissolved solids in percent. Concentration of chloramines in the water. In milligrams of sulphate per litre. S/cm measures water's electrical conductivity. In ppm, organic carbon is the percentage of organic carbon. Table of Trihalomethanes in g/L: Trihalomethanes Turbidity: In NTU, a measurement of water's light-emitting characteristic. Whether or not water is safe for human consumption is determined by its portability. Potable -1 and Non-potable -0.

3.3 The design of the proposed system:

First, we need to do image masking by taking the help of the global map. Using the VGG image annotator tool we make points around the water we found. n this way we train nearly 50 images and we store that data in a JSON file such that whenever we give a new image machine can identify it.

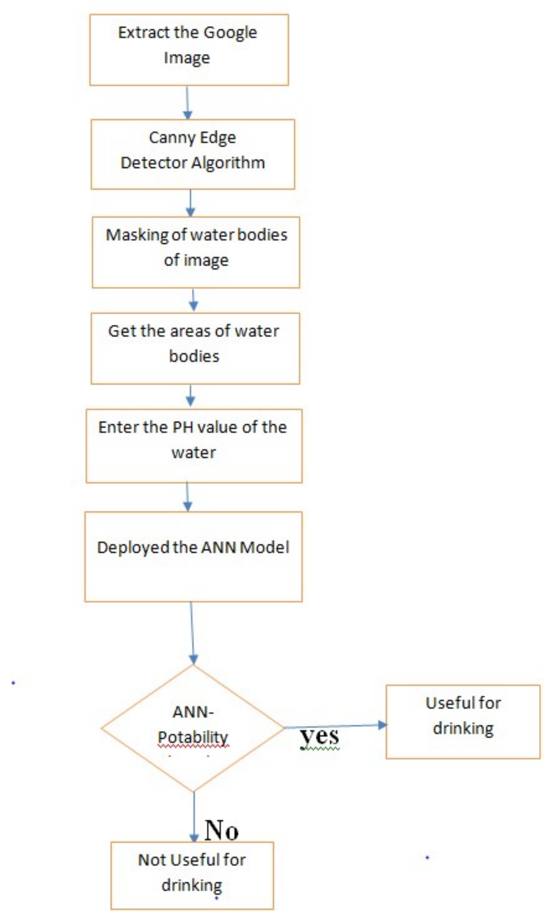

Figure.3.1: The Architecture Of Water Portability System

Journal of Theoretical and Applied Information Technology
15th October 2022. Vol.100. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

After that, we need to take a dataset that contains information about the water quality of that masked images. And finally, we need to check whether that water is suitable for drinking or not. If we get portability as 1 it is safe for drinking if we get 0 it is not safe for drinking.

3.4 Analyzing the leakage position:

Leakage location is also estimated using the below formulas:

$$d1 = \frac{1}{2} * (D + c\Delta t) ------(1)$$

$$d2 = \frac{1}{2} * (D - c\Delta t) -----(2)$$

The Smart Meter Nodes J and K, the length of the pipe, the flow rate C, and the time at which the burst occurred are all shown in the table below.
There is a link between the signals tj and tk (Sj,Sk)

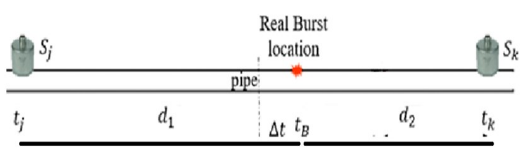

Figure.3.2: Real Burst Location Estimation Diagram

Cross-Correlation:In order to compare two signals, the signal similarity statistic is used. A signal's functional inverse is convolutional with the signal's primary signal in a correlation analysis. The cross-correlation of the two input signals is the name given to the resulting signal. The cross-correlation of x (t) and y (t) signals may be represented using a mathematical formula.

$$Rxy(\tau) = \int X(t) * Y(t - \tau)dt \quad \text{Where t range in } [-\infty, \infty] ----(1)$$

Similarly, the cross-correlation of the discrete-time signals x [n] and y [n] is expressed as

$$Rxy[m] = \sum x[n].y*[n-m] \text{ where n range in } [-\infty, \infty] ------------(2)$$

During the sample calculation, one signal is slid on top of the other while the shifting parameter, m, is established. In the case of digital signals, one signal is moved one sample to the right each time, after which the sum of the overlapping samples is determined. When the two signals being compared are the most similar, cross-correlation is at its strongest. Time and frequency are inextricably linked. (TF=1). If f equals 100, then t=1/f equals 1/100=0.01.

X1(t)= Asinωt+φ1 and X2(t)= Bsinωt+φ2, there is no phase shift, so φ1= φ2=0 and A=B=1 (Amplitude).

3.5 Algorithm to find out the time delay between two discrete time signals with different frequencies:

1) U and V are the two signals
U=sin(2*π*f1*t) and V= sin(2*π*f2*t)
2) Cal the subroutine c=XCorr(U,V), Where U and V are the two discrete-time signals with different frequencies.
3) Call the subroutine max (c) the two return values of the function are [Cmax,ICmax] Where c is the cross-correlation output, Cmax maximum value of c on the Y-axis, and Imax is the index value of Cmax in the X-axis
4) Calculate lagging time delay tlag= -Dt*(ICmax-1), Where Dt is the sampling period interval.

Where tlag is the lagging time delay of signal.

Subroutine XCorr(U,V)

Begin

1. Find the dissimilarity between these signals of discrete coordinates using Euclidian distance.
2. The product of the overlapping samples is determined each time a signal is pushed to the right by a single sample.
3. When the two signals under consideration become the most similar to each other, the cross-correlation achieves its maximum.
4. Return this max value of cross correlation value as c.

End

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Subroutine max(c)

Begin

1. Maxx1= first signal index values of X-axis, when it reaches to Peek.
2. Maxx2= Index value of (C)
3. return (Maxx1-Maxx2)

End

4. RESULTS & DISCUSSION

After masking we will collect a dataset of that water bodies, which contains various features such as PH, Hardness, solids, chloramines, organic _ carbon, sulphates,trilohamithines, turbity, and potability. Now we need to check for portability using all the inputs available if portability is 1 then it is safe for drinking else it is not safe for drinking.

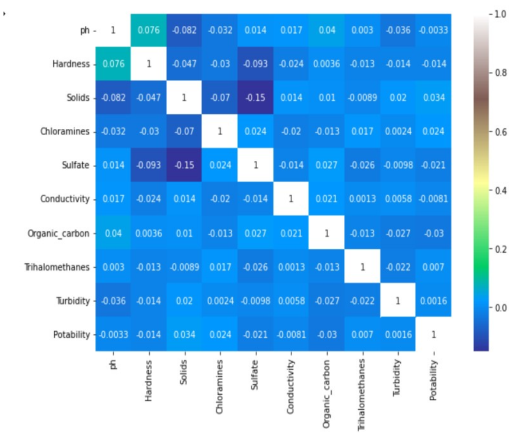

Figure.4.1: Confusion Matrix For The Water Portability Features

Plot a heat map that shows the correlation between each output to remove any attribute their correlation should be greater than 75% but we are not getting that so we are not removing any attribute.

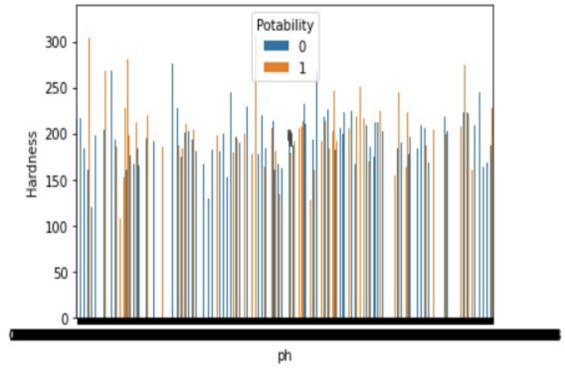

Figure.4.2: Bar Chart For The Water Portability Features Hardness To Ph

Bar chart representation of various Modeling techniques:The various modeling are decision tree, decision tree with hyper parameter tuning, logistic regression, logistic regression with hyper parameter tuning, k nearest neighbor, and artificial neural networks are shown in figure 3.2

In the below diagram (Fig 4.4) one hundred two-dimensional coordinates are identified, and then measure the distance between these two sine waves, using the Euclidian distance formula by adopting the leakage localization. Algorithm [17].

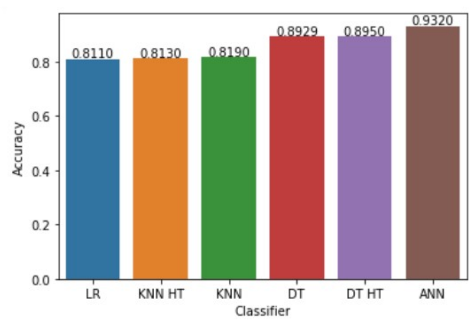

Figure.4.3: The Various Classification Modeling Techniques With Their Accuracy

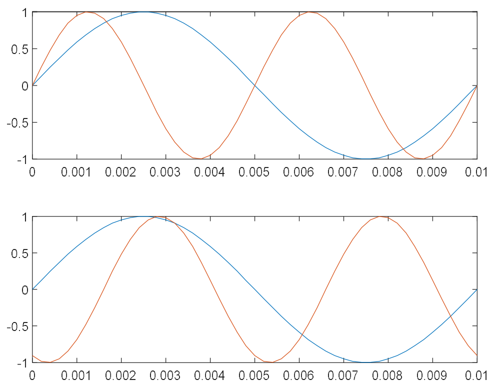

Figure.4.4: Measuring The Time Delay Between Two Sinusoidal Signals With Different Frequencies.

The first signal indicated with magenta color, it has collected from smart meter Sj, making 200 cycles per unit, then it makes two cycles per 0.01 units. The second signal which is indicated with blue color was collected from smart water meter node Sk, making 100 cycles per unit time. It indicates there is a leakage in the pipeline; it reduces half of the existing cycles. The sampling period interval here is considered as 0.0002 units. At each time the first signal shifted one position right and then evaluates the dot product of two signals. Where this dot product value is max, then there is the max similarity between these two signals. The above algorithm illustrated this technique. The second diagram shows the cross-correlated values of these two signals in the two-dimensional plane.

This bar chart represents (Fig 4.5) the real burst location from the smart water meter nodes, which locations are identified by the longitude and latitude features fromthe original dataset. meter node Sj and d2 distance from the smart meter Sk.

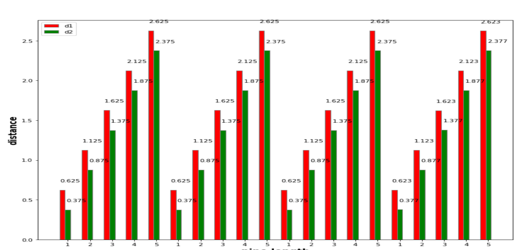

Figure.4.5: Bar Chart Between The Pipe Length And Distances D1 And D2.

It indicates that burst is d1 distance from smart water Irrespective of any variations in the flow rate and pulse rate are not shown to influence the real burst location.

5. CONCLUSION

Using Google Earth images of water bodies, we can identify water bodies and check their quality. Check the quality of water is reaching the specified portability range or not, if it has not satisfied the portability constraints then it can be used for domestic purposes. For the feature enhancement, there is provision to integrate our project with Django web framework flask and create a website such that when the user gives Google Earth images of water bodies then the website does mask and also gives the portability of water whether it is safe or not for drinking purpose. The leakage positions are predicted and represented in the bar chart for various pipe lengths and exact leakage in the water distribution pipelines.

REFERENCES:

[1] S. Thenmozhi, K. Sumathi, A. Asokan, B. Priyanka, R. Maheswar and P. Jayarajan, "IoT Based Smart Water Leak Detection System for a Sustainable Future," 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2021, pp. 359-362, doi: 10.1109/WiSPNET51692.2021.9419456.

[2] M. T. Islam and S. Aslan, "Leak Detection and Location Pinpointing in Water Pipeline Systems Using a Wireless Sensor Network," 2021 IEEE International Symposium on Circuits and Systems (ISCAS), 2021, pp. 1-7, doi: 10.1109/ISCAS51556.2021.9401106.

[3] H. Lin, H. Lin, X. Fang, M. Wang and L. Huang, "Intelligent Pipeline Leak Detection and Analysis System," 2020 15th International Conference on Computer Science & Education (ICCSE), 2020, pp. 206-210, doi: 10.1109/ICCSE49874.2020.9201761.

[4] S. Porwal, S. A. Akbar and S. C. Jain, "Leakage detection and prediction of location in a smart water grid using SVM classification," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017, pp. 3288-3292, doi: 10.1109/ICECDS.2017.8390067.

[5] Roy, U. “Leak Detection in Pipe Networks Using Hybrid ANN Method”. Water Conserv Sci Eng 2, 145–152 (2017). https://doi.org/10.1007/s41101-017-0035-1

[6] Sena, P.. (2013). An Optimal Heuristic for Sum of All Prime Numbers Logic for Large Inputs using RAPTOR.

[7] Sena, P. Vasanth, and P. Sammulal. "A Survey on Leaks and Faults Detection & Diagnosis in Water Supply System." International Journal of Engineering, Science and Mathematics 7.2 (2018): 174-180.

[8] Nune, Ganesh Kumar, and P. Vasanth Sena. "Novel artificial neural networks and logistic approach for detecting credit card deceit." International Journal of Computer Science and Network Security (IJCSNS) 15.9 (2015): 21.

[9] Sena, P. Vasanth, Sammulal Porika, and M. Venu Gopalachari. "A Mining Framework for Efficient Leakage Detection and Diagnosis in Water Supply System." ICCCE 2020. Springer, Singapore, 2021. 1093-1103.

[10] Sena, P. Vasanth, and Porika Sammulal. "An Optimal Heuristic for Student Failure Detection and Diagnosis in the Sathvahana Educational Community Using WEKA." International Conference on Communications and Cyber Physical Engineering 2018. Springer, Singapore, 2018.

[11] WojciechMoczulski, "The case study - A methodology of leakage detection and location in water distribution networks", 3rd Conference on Control and Fault-Tolerant Systems, IEEE 2016.

[12] Abdullah Kadri, "Characterization of an acoustic wireless sensor for water leakage detection in underground pipes", Elsevier, 2012.

[13] Bansal, Anshul, Susheel Kaushik Rompikuntla, Jaganadh Gopinadhan, Amanpreet Kaur and Zahoor Ahamed Kazi. "Energy Consumption Forecasting for Smart Meters." ArXiv abs/1512.05979, Cornell university, Elsevier, 2015.

[14] Misiunas, "Failure monitoring and asset condition assessment in water supply system", PhD Thesis, Department of Electrical Engineering and Automation, Lund University, Sweden,2005.

[15] Feng, J., Zhang, H., & Liu, "Application of Fuzzy Decision-Making in Pipeline Leak Localization. Budapest", Hungary: IEEE. 25–29, 2004.

[16] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 315–323, 2011.

[17] Sena, P. Vasanth, and P. Sammulal. "A Survey on Leaks and Faults Detection & Diagnosis in Water Supply System." International Journal of Engineering, Science and Mathematics 7.2 (2018): 174-180.

[18] Nune, G. K., & Sena, P. V. (2015). Novel artificial neural networks and logistic approach for detecting credit card deceit. International Journal of Computer Science and Network Security (IJCSNS), 15(9), 21.

[19] Vasanth Sena, P., Porika, S., Venu Gopalachari, M. (2021). A Mining Framework for Efficient Leakage Detection and Diagnosis in Water Supply System. In: Kumar, A., Mozar, S. (eds) ICCCE 2020. Lecture Notes in Electrical Engineering, vol 698. Springer, Singapore. https://doi.org/10.1007/978-981-15-7961-5_101

[20] Vasanth Sena, P., Sammulal, P. (2019). An Optimal Heuristic for Student Failure Detection and Diagnosis in the Sathvahana Educational Community Using WEKA. In: Kumar, A., Mozar, S. (eds) ICCCE 2018. ICCCE 2018. Lecture Notes in Electrical Engineering, vol 500. Springer, Singapore. https://doi.org/10.1007/978-981-13-0212-1_68

[21] M Venu Gopalachari, Morarjee Kolla, Rupesh Kumar Mishra, Zarin Tasneem, "Design and Implementation of Brain Tumor Segmentation and Detection Using a Novel Woelfel Filter and Morphological Segmentation", Complexity, vol. 2022, Article ID 6985927, 9 pages, 2022.

[22] S. Rakesh, Nagaratna P Hegde, "A Framework for Intelligent Traffic Control System" International Journal of Engineering Trends and Technology, vol. 70, no. 7, pp. 88-93, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I7P210