Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 10, July-December: 2022

WATER QUALITY PREDICTION AND MONITORING USING AUTO REGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

Dr. BTR Naresh Reddy[1] , A. Mallesh[2] , G. Archana Reddy[3]

[1] Professor, Department of Computer Applications,Aurora's PG College (MBA), Uppal, Hyderabad
Email: reddynareshbtr@gmail.com

[2] Assistant Professor, Department of Computer Applications,Aurora's PG College (MBA), Uppal, Hyderabad
Email: mallesh67@gmail.com

[3] Assistant Professor, Department of Computer Applications,Aurora's PG College (MBA), Uppal, Hyderabad
Email: archanareddy054@gmail.com

ABSTRACT

One of the most important resources in daily living is water. Source water is now more vulnerable to contamination due to pollution and urbanization. The development of a water quality monitoring system is necessary in order to maintain the quality of source water and ultimately protect public health. This research suggests a low-cost wireless water quality monitoring system that continuously checks the condition of water kept in above tanks. Critical quality indicators are used to measure the water's quality. These parameters' worth of data, together with their timestamp, are kept in real-time in a cloud database. The observed data is compared to standard, set thresholds in order to determine the water quality. The timestamp-annotated data is regarded as a time series. To forecast individual water quality indices, a univariate non-seasonal Auto Regressive Integrated Moving Average (ARIMA) model is used. Water quality degradation is forecasted using forecasting findings. The mean square errors of the model between the actual and predicted values are determined to be 0.001 for turbidity, 0.076 for temperature, and 0.001 for pH.

Keywords: Water Quality, Prediction, Monitoring, Auto Regressive Integrated Moving Average (ARIMA)

1. INTRODUCTION

Above-ground storage tanks are frequently utilized to store water for residential or commercial uses. Numerous infections or dangerous bacteria may establish a breeding habitat in the water that is being kept. Rainwater loses its acidity when it comes into contact with it, making it unfit for drinking and other uses. Over time, toxic substances could accumulate on the tank walls. Particulate matter pollution may result from outdoor exposure. These particles can sediment and change the water's chemical composition. The rust that forms on water collection pipes as a result of poor maintenance drastically reduces water quality. The microbiological quality of water is a sign of illness. Water contamination is a common way for infectious diseases like giardiasis, cholera, typhoid, guinea worm, hepatitis, and schistosomiasis to spread. Inadequate hygiene or sanitation practices may be the cause of

several illnesses. It is crucial to remember that every element pertaining to the availability and quality of drinking water matters and is connected to public health. Prior to addressing these issues, a real-time system that continuously records and reports water quality is suggested. The system's data collection accurately depicts the quality of the water. Predicting the decline in water quality can be done with the help of a helpful examination of this data. Using time-series forecasting will help achieve this. A historical time series is analyzed using the forecasting statistical technique known as Auto Regressive Integrated Moving Average (ARIMA). Since the model needs to be immune to any short-lived local trends within the time series, a non-seasonal model is a good fit. The general quality of the water in the future is not improved by these tendencies.

2. LITERATURE SURVEY AND RELATED WORK

Physicochemical parameters for testing of water-A review:
Due to increase population, advanced agricultural practices, industrialization, man- made activity, water is being highly polluted with different contaminants. Water is a vital resource for human survival. The availability of good quality water is an indispensable feature for preventing diseases and improving quality of life. It is necessary to know details about different physico-chemical parameters such as colour, temperature, Total hardness, pH, sulphate, chloride, DO, BOD, COD, alkalinity used for testing of water quality.

Modeling, development & analysis of low cost device for water quality testing
In recent days, the most important problem that our society faces is low quality of drinking water. Water quality monitoring is important because contaminated drinking water can spread diseases faster than any other sources. With existing techniques, the general public is not aware of the potability of water. Lack of accurate and efficient low cost systems are a reason for poor awareness on the same. This paper focuses on modelling and developing a low cost water quality testing device and analysing its performance with the currently available products. The developed device can measure the parameters like pH, Total Dissolved Solids, Conductivity and Temperature. Its results are verified with samples of distilled water, salt water, tap water, dish wash and water, curd and performance is studied
Implementation of wireless sensor network for real time overhead tank water quality monitoring
Water is a precious source vital for healthy living. Most of the infectious diseases are due to contaminated water which leads to millions of deaths every year. There is a need to establish Water quality monitoring system to verify whether the determined water quality is suitable for intended use. This paper presents the application of Wireless Sensor Network (WSN) technology for real time online Water quality monitoring. In this paper, the details of system design and implementation of WSN are presented. Wireless Sensor Network (WSN) for a water quality monitoring is composed of number of sensor nodes with networking capability which are deployed at different overhead tanks and water bodies in an area. Each sensor node consists of an Arduino microcontroller, Xbee module and water quality sensors, the sensor probes shall continuously measure the different water quality parameters like pH, Temperature, Conductivity. The parameters are measured in real time by the sensors and send the data to the data center. Solar panel is used to power the system for each node. Data collected from remote nodes are displayed in the user PC. This developed system will demonstrate online sensor data analysis and has the advantages of power optimization, portability and easy installation

Water quality monitoring system using zigbee based wireless sensor network:
Water is an essential natural resource that is fundamental to both human and animal health and more than 88% of South African households have access to water. Although South Africa has one of the cleanest water systems globally but there are still many factors causing water pollution. Industries produce a lot of waste that can affect the water PH, sometimes the water reservoir is exposed to dust, the Agric sector causes soil erosion due to its physical soil irritation or not enough care is taken by the departmental authorities. Although there's effort in urban areas by municipalities to purify and monitor the water used for consumption, the existing mechanism used is not as automated and reliable, even worse in rural areas no proper care is taken or no genuine mechanisms put in place. In this paper, Water Quality with Notifications System is presented to better the lives of the people by providing an easy and simple mechanism to know the water quality stored in the reservoir. It holds different sensors for getting water quality parameters namely PH Sensor, Nitrate Sensor, Turbidity Sensor, Water Level Detection Sensor. The system gets water parameters from the sensors through Raspberry Pi. It is a portable system and no monitoring authority is required at the water source 24 hours a day. Also, it is a mobile and portable system that is connected to an android smart phone.

3. EXISTING SYSTEM

The qualitative characteristics for which the system is designed are temperature, pH, turbidity and total dissolved salts (TDS). The system consists of sensors that serve the purpose of data collection, a core for processing, and an RF module that transmits the processed data to a master receiver. An Arduino microcontroller forms the core. The RF module used is a standard 433MHz module. The receiver sends the processed data to the cloud. The above system is designed for the transmitter to be placed at the over-head tank and a receiver to be placed anywhere within a 50-yard radius (from the transmitter.

4. PROPOSED SYSTEM

the proposed framework has settled to Water Quality Monitoring framework which depends on Machine to Machine correspondence through AI. Turbidity and the conductivity sensors are connected. The conductivity acts as a sensor gateway. The sensor input are sent to the pi4, a edge level processor(personal computer) where in the K Means, a machine learning algorithm is used for predicting the quality of water.The predicted water quality data are stored in Cloud server for future access. The predicted data is sent to the water controller unit for further action.This has brought about complete computerized Water Quality Monitoring framework utilizing IoT and AI Technologies by which the gadgets impart among themselves in anticipating the Water Quality for private country region. Subsequently, the water quality can be watched consequently with no human interference. The proposed system can be extended further by water retreatment mechanism

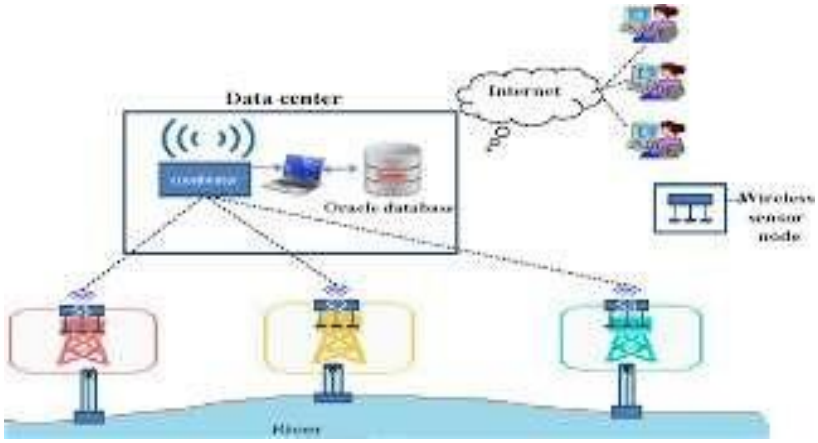


FIG 1- SYSTEM ARCHITECTURE

5. IMPLEMENTATION

Modules Used in Project :-

Tensorflow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.

TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

It is the fundamental package for scientific computing with Python. It contains various features including these important ones:

- A powerful N-dimensional array object
- Sophisticated (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using Numpy which allows Numpy to seamlessly and speedily integrate with a wide variety of databases.

Pandas

Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter Notebook, web application

servers, and four graphical user interface toolkits. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatter plots, etc., with just a few lines of code. For examples, see the sample plots and thumbnail gallery.

For simple plotting the pyplot module provides a MATLAB-like interface, particularly when combined with IPython. For the power user, you have full control of line styles, font properties, axes properties, etc, via an object oriented interface or via a set of functions familiar to MATLAB users.

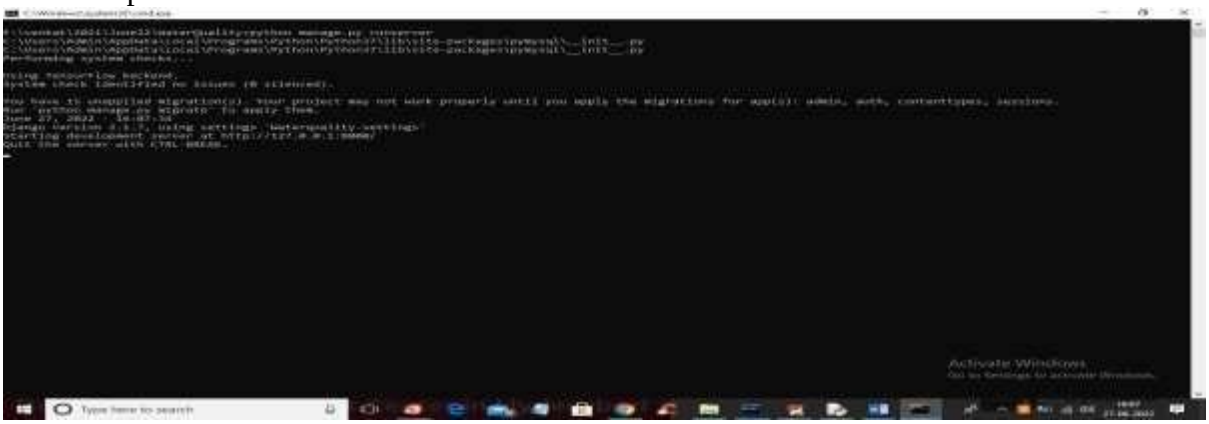
Scikit – learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

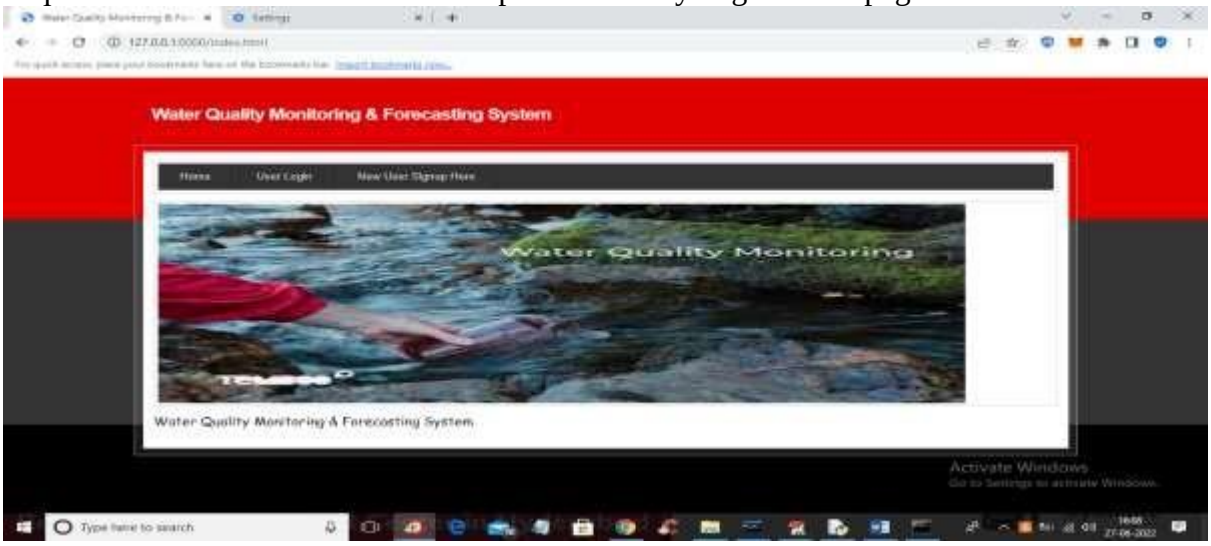
5. RESULTS AND DISCUSSION SCREENSHOTS

SCREEN SHORTS:

To run project first copy content from 'DB.txt' file and then paste in MYSQL database to create it and now double click on 'run.bat' file to start DJANGO server and then will get below output

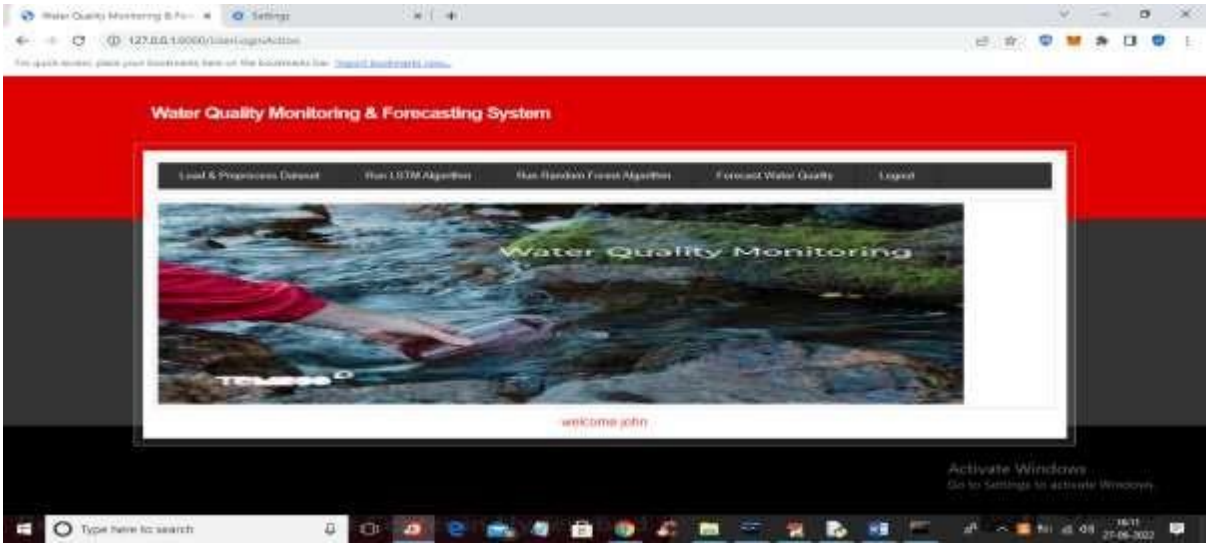


In above screen DJANGO server started and now open browser and enter URL as 'http://127.0.0.1:8000/index.html' and press enter key to get below page



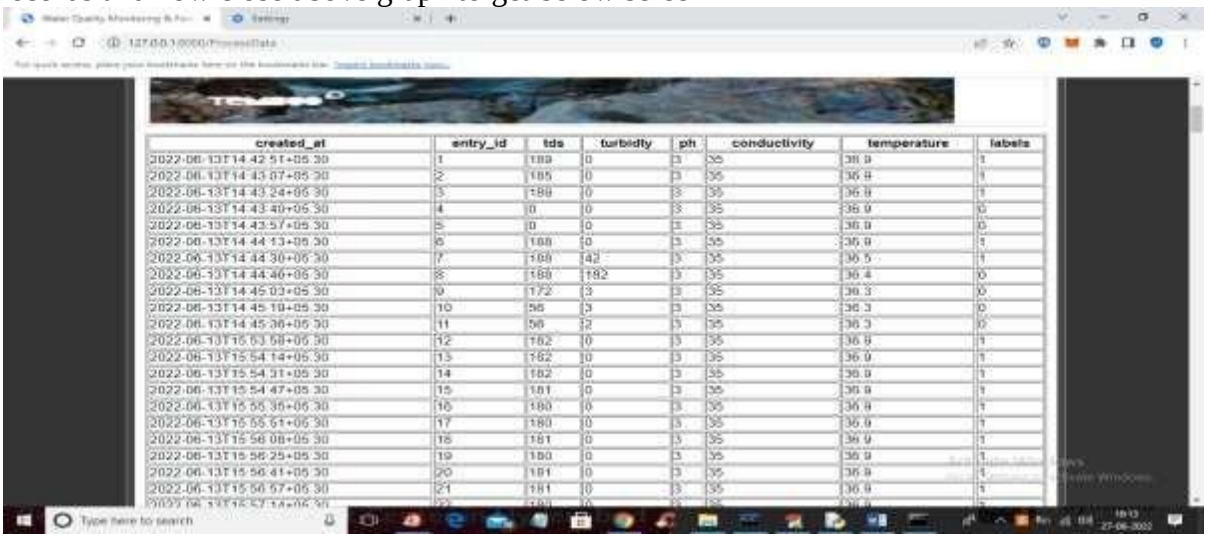
In above screen click on 'New User Signup Here' link to get below screen

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 10, July-December: 2022



In above screen click on 'Load & Preprocess Dataset' link to load and process dataset such as replacing missing values with 0 and then split dataset into train and test and get below output

In above screen dataset is processed and in above graph x-axis contains water quality as 0 or 1 where 0 means GOOD quality and 1 means POOR quality and y-axis represents number of records and now close above graph to get below screen

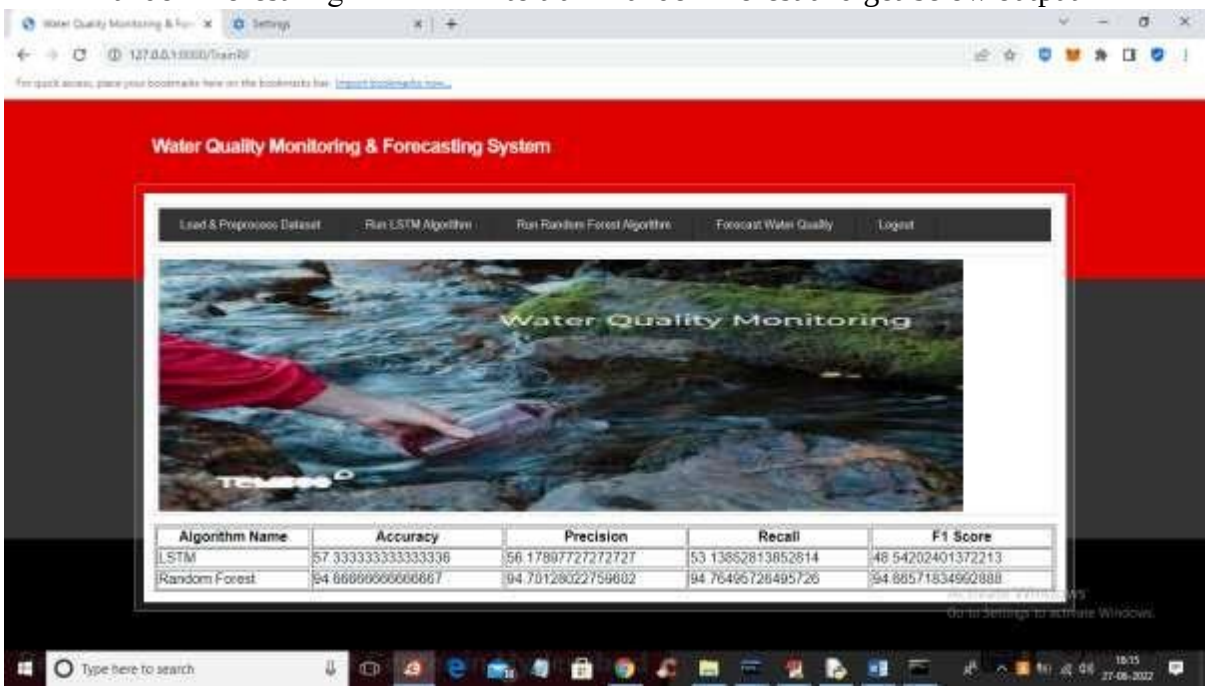


created_at	entry_id	tds	turbidity	ph	conductivity	temperature	labels
2022-06-13T14:42:51+05:30	1	189	0	3	35	36.9	1
2022-06-13T14:43:07+05:30	2	185	0	3	35	36.8	1
2022-06-13T14:43:24+05:30	3	189	0	3	35	36.8	1
2022-06-13T14:43:40+05:30	4	0	0	3	35	36.9	0
2022-06-13T14:43:57+05:30	5	0	0	3	35	36.8	0
2022-06-13T14:44:13+05:30	6	188	0	3	35	36.9	1
2022-06-13T14:44:30+05:30	7	188	142	3	35	36.5	1
2022-06-13T14:44:46+05:30	8	189	182	3	35	36.4	0
2022-06-13T14:45:03+05:30	9	172	3	3	35	36.3	0
2022-06-13T14:45:19+05:30	10	56	3	3	35	36.3	0
2022-06-13T14:45:36+05:30	11	56	2	3	35	36.3	0
2022-06-13T15:53:58+05:30	12	182	0	3	35	36.8	1
2022-06-13T15:54:14+05:30	13	182	0	3	35	36.9	1
2022-06-13T15:54:31+05:30	14	182	0	3	35	36.8	1
2022-06-13T15:54:47+05:30	15	181	0	3	35	36.8	1
2022-06-13T15:55:35+05:30	16	180	0	3	35	36.8	1
2022-06-13T15:55:51+05:30	17	180	0	3	35	36.9	1
2022-06-13T15:56:08+05:30	18	181	0	3	35	36.9	1
2022-06-13T15:56:25+05:30	19	180	0	3	35	36.9	1
2022-06-13T15:56:41+05:30	20	181	0	3	35	36.8	1
2022-06-13T15:56:57+05:30	21	181	0	3	35	36.8	1

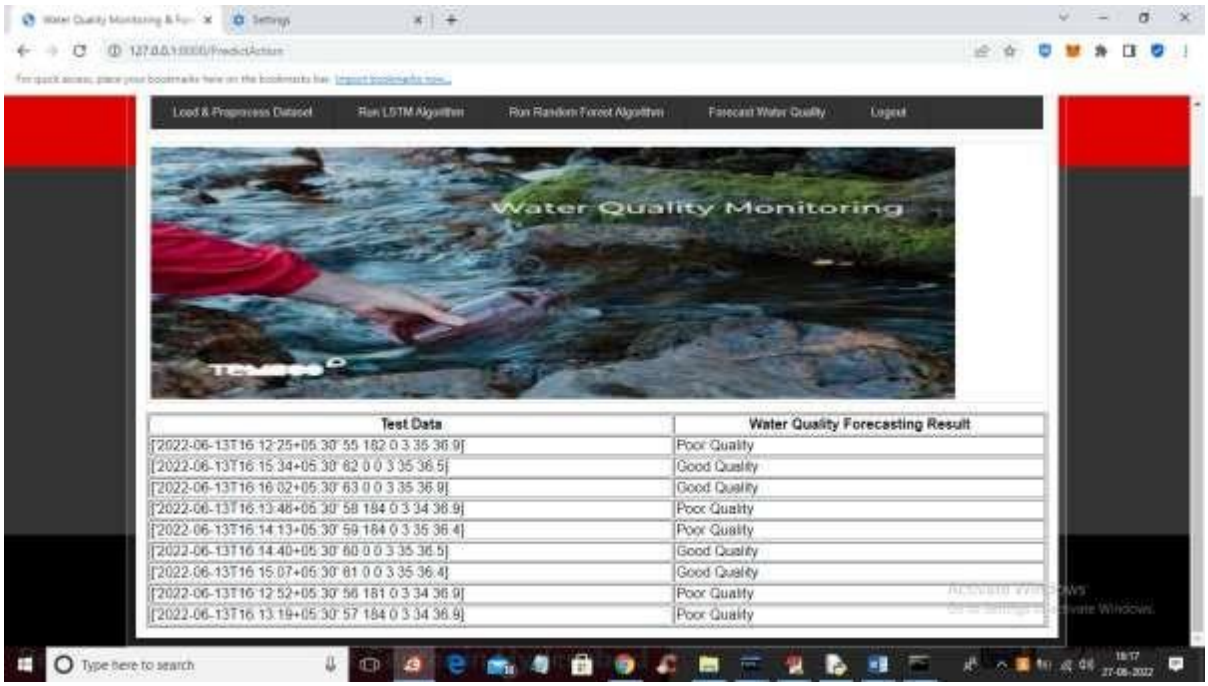
In above screen we can see dataset processed and loaded and now click on 'Train LSTM Algorithm' link to train LSTM and get below output

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 10, July-December: 2022

In above screen LSTM got trained and with LSTM we got 57% accuracy and now click on 'Train Random Forest Algorithm' link to train Random Forest and get below output



In above screen with Random Forest we got 94% accuracy and now click on 'Forecast Water Quality' link to upload test data and then forecast quality



In above screen in tabular output first column contains water test values and second column contains forecast result as 'Poor' or "Good"

7. CONCLUSION AND FUTURE SCOPE

In this paper, a wireless, efficient, low power, low-cost, stable monitoring system is proposed to perform real-time water quality monitoring and reporting. The error-free data is collected and stored in the cloud database.

REFERENCES

[1] Patil, P. N., D. V. Sawant, and R. N. Deshmukh. "Physicochemical parameters for testing of water-A review." International Journal of Environmental Sciences 3.3 (2012): 1194.

[2] Indu, K., and Jishmi Jos Choondal. "Modeling, development & analysis of low cost device for water quality testing." 2016 IEEE Annual India Conference (INDICON). IEEE, 2016.

[3] Sowmya, Ch, et al. "Implementation of wireless sensor network for real time overhead tank water quality monitoring." 2017 IEEE 7th International Advance Computing Conference (IACC). IEEE, 2017.

[4] Rasin, Zulhani, and Mohd Rizal Abdullah. "Water quality monitoring system using zigbee based wireless sensor network." International Journal of Engineering & Technology 9.10 (2009): 24-28

[5] Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. Time Series Analysis: Forecasting and Control;John Wiley & Sons: Hoboken, NJ, USA, 2015.

[6] Tyralis, H.; Papacharalampous, G. Variable selection in time series forecasting using random forest algorithms 2017,10, 114.

[7] Hernández, Nathalie, et al. "Arima as a forecasting tool for water quality time series measured with UV-Vis spectrometers in a constructed wetland." Tecnología y ciencias del agua 8.5 (2017): 127-139.

[8] An, Qi, and Min Zhao. "Time Series Analysis in the Prediction of Water Quality." 7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017). Atlantis Press, 2017.