**LITERATURE SURVEY**

**1.Ahmed U, Mumtaz R, Anwar H et al (2019) “Efficient water quality prediction using supervised machine learning”**.

Water makes up about 70% of the earth’s surface and is one of the most important sources vital to sustaining life. Rapid urbanization and industrialization have led to a deterioration of water quality at an alarming rate, resulting in harrowing diseases. Water quality has been conventionally estimated through expensive and time-consuming lab and statistical analyses, which render the contemporary notion of real-time monitoring moot. The alarming consequences of poor water quality necessitate an alternative method, which is quicker and inexpensive. With this motivation, this research explores a series of supervised machine learning algorithms to estimate the water quality index (WQI), which is a singular index to describe the general quality of water, and the water quality class (WQC), which is a distinctive class defined on the basis of the WQI. The proposed methodology employs four input parameters, namely, temperature, turbidity, pH and total dissolved solids. Of all the employed algorithms, gradient boosting, with a learning rate of 0.1 and polynomial regression, with a degree of 2, predict the WQI most efficiently, having a mean absolute error (MAE) of 1.9642 and 2.7273, respectively. Whereas multi-layer perceptron (MLP), with a configuration of (3, 7), classifies the WQC most efficiently, with an accuracy of 0.8507. The proposed methodology achieves reasonable accuracy using a minimal number of parameters to validate the possibility of its use in real time water quality detection systems.

**ADVANTAGES**

Solve problems related to water quality Local research employing lab analysis helped us gain a greater insight into the water quality problem in Pakistan. Previous studies used more than 10 water quality parameters to predict WQI whereas, our methodology employs only four water quality parameters to predict WQI, with a MAE of 1.96, and to predict water quality class with an accuracy of 85%.

**2. Aldhyani THH, Al-Yaari M, Alkahtani H, Maashi M (2020) “Water quality prediction using artificial intelligence algorithms”.**

During the last years, water quality has been threatened by various pollutants. Therefore, modeling and predicting water quality have become very important in controlling water pollution. In this work, advanced artificial intelligence (AI) algorithms are developed to predict water quality index (WQI) and water quality classification (WQC). For the WQI prediction, artificial neural network models, namely nonlinear autoregressive neural network (NARNET) and long short-term memory (LSTM) deep learning algorithm, have been developed. In addition, three machine learning algorithms, namely, support vector machine (SVM), -nearest neighbor (K-NN), and Naive Bayes, have been used for the WQC forecasting. The used dataset has 7 significant parameters, and the developed models were evaluated based on some statistical parameters. The results revealed that the proposed models can accurately predict WQI and classify the water quality according to superior robustness. Prediction results demonstrated that the NARNET model performed slightly better than the LSTM for the prediction of the WQI values and the SVM algorithm has achieved the highest accuracy (97.01%) for the WQC prediction. Furthermore, the NARNET and LSTM models have achieved similar accuracy for the testing phase with a slight difference in the regression coefficient ( and ). This kind of promising research can contribute significantly to water management.

**ADVANTAGES**

Modeling and prediction of water quality are very important for the protection of the environment. Developing a model by using advanced artificial intelligence algorithms can be used to measure the future water quality. In this proposed methodology, the advanced artificial intelligence algorithms, namely, NARNET and LSTM models were used to predict the WQI. Moreover, machine learning algorithms such as SVM, KNN, and Naive Bayes were used to classify the WQI data. The proposed models were evaluated and examined by some statistical parameters. For the WQI prediction, the result has revealed that the performance of the NARNET model is slightly better than the LSTM model based on the obtained value. However, the SVM algorithm has achieved the highest accuracy of the prediction of the WQC as compared with KNN and Naive Bayes algorithms. After examining the robustness and efficiency of the proposed model for predicting the WQI, in future work, the developed models will be implemented to predict the water quality in Saudi Arabia for different types of water.

**3.Asadollah SBHS, Sharafati A, Motta D, Yaseen ZM (2021) “River water quality index prediction and uncertainty analysis”.**

The Water Quality Index (WQI) is the most common indicator to characterize surface water quality. This study introduces a new ensemble machine learning model called Extra Tree Regression (ETR) for predicting monthly WQI values at the Lam Tsuen River in Hong Kong. The ETR model performance is compared with that of the classic standalone models, Support Vector Regression (SVR) and Decision Tree Regression (DTR). The monthly input water quality data including Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), Electrical Conductivity (EC), Nitrate-Nitrogen ( -N), Nitrite-Nitrogen ( -N), Phosphate (), potential for Hydrogen (pH), Temperature (T) and Turbidity (TUR) are used for building the prediction models. Various input data combinations are investigated and assessed in terms of prediction performance, using numerical indices and graphical comparisons. The analysis shows that the ETR model generally produces more accurate WQI predictions for both training and testing phases. Although including all the ten input variables achieves the highest prediction performance (, ), a combination of input parameters including only BOD, Turbidity and Phosphate concentration provides the second highest prediction accuracy (, ). The uncertainty analysis relative to model structure and input parameters highlights a higher sensitivity of the prediction results to the former. In general, the ETR model represents an improvement on previous approaches for WQI prediction, in terms of prediction performance and reduction in the number of input parameters.

**ADVANTAGES**

A new ensemble machine learning model is developed to predict the Water Quality Index (WQI).Observed water quality variables in the Lam Tsuen River in Hong-Kong are used to predict the WQI.The prediction uncertainty associated with model structure and input variable selection is quantified. The three modeled considered, and the ETR model in particular, provides accurate WQI predictions.

**4.Dezfooli D, Hosseini-Moghari S-M, Ebrahimi K, Araghinejad S (2018) “Classification of water quality status based on minimum quality parameters”.**

This paper focuses on three models namely probabilistic neural network (PNN), k-nearest neighbor and support vector machine (SVM) as an alternative to NSFWQI in order to classify water quality of Karoon River, Iran as a case study, regarding minimum possible parameters. For this purpose a set of 172 water samples were used in a way that water quality parameters and their water quality classes (from NSFWQI) were considered as the input–output of the models, respectively. Three assessment criteria, namely error rate, error value and accuracy, were applied in order to assess the performance of the applied models. The results revealed that under the condition that no parameter is removed, all the three models showed the same results. However, under quality parameters removal the results revealed that PNN was the best model, as that managed to classify water quality solely based on three quality parameters of turbidity, fecal coliform and total solids, at 90.70% accuracy, 9.30% error rate and error value was 4. In contrast to PNN, in the same condition, SVM showed the poorest performance. As with making use of four quality parameters namely fecal coliform, DO, BOD and total phosphates, it classified water quality at 88.37% accuracy and 11.63% error rate.

**ADVANTAGES**

Given the fact that water quality classiﬁcation by means of water quality indices is time-con-summing and costly, data-driven methods were proposed in the present study as an alternative to classify water quality. The results obtained from NSFWQI revealed that samples could be classiﬁed into three categories of good, bad and medium. Making use of three water quality parameters of turbidity, fecal coliform and total solids, PNN model with the accuracy of 94.57 and 90.70% at the calibration and testing phases respectively delivered the most reasonable performance compared to other two models. PNN model may, therefore, reduce the sampling costs and computation time for water quality classiﬁcation and it could be regarded an appropriate alternative to NSFWQI. Moreover, the results obtained from the comparisons made among the three models revealed that fecal coliform is the most eﬀective parameter in water quality classiﬁcation since if omitted, a great error might occur in water quality classiﬁcation. It is recommended that other models be employed for water quality classiﬁcation in future studies in order to determine the water quality parameters which might have an eﬀect on water quality classiﬁcation.