Explainable AI Framework for Multi-label Classification using Supervised Machine Learning Models

Siddarth **Singaravel[1]**, Dibyanshu **Jaiswal[1]**, Jay **Daftari[1]**, Shridevi S [2]

[1]*School of Computer Science and Engineering, Vellore Institute of Technology, India*
[2]*Centre for Advanced Data Science, Vellore Institute of Technology, India*

Abstract
Water is known as a "universal solvent" as it is extraordinarily frail against contamination. Water quality standards are developed based on logical evidence on the effects of hazardous compounds on a certain quantity of water used. Classification technique of machine learning can be employed to under-stand the water quality status. In this work, supervised machine learning models are being implemented to classify water quality indexes, and the Smote analysis is used to handle the imbalance in the dataset. Artificial neural net-work model is built using the features such as Oxygen, pH, temperature, total suspended sediment, turbidity, nitrogen, and phosphorus as inputs and water quality check as target variable. This target variable is created using Canadian Council of Ministers of the Environment Water Quality Index, and the model works with an accuracy of 87%. The classification is done on XGBoost model as well and it performs with an accuracy of 90%. The explanations for predictions of these models for a data instance were performed using explainable artificial intelligence tools such as LIME and SHAP. The results and interpretations for the predictions seem to be more promising and attractive making the proposed models more interpretable, accurate and efficient. Through our re-search we can benefit our readers by providing them clarity about exactly what features are having more influence on water quality than others from different machine learning algorithms. This will help the developers to gain insights about the significant factors of poor water quality and how to overcome that.

Keywords:
Lime, SHAP, Explainable AI, Artificial neural network, Super-vised machine learning, gradient boosting, multi-label classification

1. Introduction

The most important factor and fundamental for supporting a wide range of life is Water. It is possibly the most transferable medium with quite a far reach. On an average human consumes 80-100 gallons of water per day. As detailed, in agricultural nations, Water-borne diseases account for 80% of the illnesses, resulting in 5 million passes and 2.5 billion illnesses [1]. That's why one has to consider the quality of water very seriously. Contaminated water causes lethal diseases like cholera, diarrhoea, amebiasis, hepatitis, gastroenteritis, etc. Unsafe water pollutants incorporate harmful substances at low concentrations. Cancer-causing, mutagenic and erotogenic can be toxic, especially when they are persevering. To lessen contributions of phosphorus, nitrogen, and pesticides from nonpoint sources (especially horticultural sources) to water bodies, natural and agrarian experts in an expanding number of nations are specifying the need to utilize best ecological practices. Some other water quality factors, like disintegrated oxygen, water quality parameters are set at the least adequate focus

ACI'22: Workshop on Advances in Computation Intelligence, its Concepts & Applications at ISIC 2022, May 17-19, Savannah, United States
EMAIL: s.siddarth2018@vitstudent.ac.in , dibyanshu.jaiswal2018@vitstudent.ac.in, jay.daftari2018@vitstudent.ac.in , shridevi.s@vit.ac.in
ORCID: 0000-0002-9877-2468; 0000-0002-7654-1570; 0000-0001-9491-6842; 0000-0002-0038-7212

to guarantee the support of natural capacities. Water quality models depend on factors that portray the nature of water and the nature of the suspended particulate matter, the base residue, and the biota. In this regard, the primary goal of this research is to propose and evaluate supervised machine learning models like ANN and XG-Boost to efficiently classify water quality. The work is carried on the dataset of "WQI Parameter Scores 1994-2013", procured by the WA State Department of Ecology's River. WQI (Water Quality Index) values are categorized into different classes based on designations set by CCME (Canadian Council of Ministers of the Environment). The proposed approach consists of 3 phases: pre-processing, feature extraction, and classification. Then the categorizing is followed by a modern method for explaining the black-box model using XAI (explainable AI) tools like LIME and SHAP. The implication and the novelty of the work on these XAI tools are mentioned.

The rest of this research paper is organized as follows. Section 2 tells about the recent research works done for predicting water quality using different machine learning models. Section 3 describes the proposed work of ANN, XG-Boost algorithms and Explainable AI. Section 4 discusses the results of discrete XAI tools and error percentage metrics for the prior mentioned models. Finally, Section 5 summarizes the findings and discusses the future research.

2. Related Work

Contaminated water can prompt some waterborne infections and impact child mortality. To decrease the impact of sullied water, it is fundamental to evaluate unique aspects of water quality. [2] Compares supervised and unsupervised learning to examine different parameters which play a crucial role in deciding water quality. They take three parameters that the dataset contains: ph., dissolved oxygen, and turbidity.

Water quality has deteriorated as a result of higher pollution concentrations. More and more big data is produced at a high rate in the process [3] of building and operating relevant water quality monitoring systems based on the Internet of Things (IoT), which has complicated water quality data. A drinking-water quality model can be recommended to anticipate water quality using sophisticated deep learning (DL) theory and long short-term memory (LSTM) performance in time-series prediction. The water quality data collected by the Guangzhou Water Source of the Yangtze River's automatic water quality monitoring station was used to investigate the parameters of water quality in greater depth, and the prediction model was trained and tested using monitoring data from January 2016 to June 2018. The study's findings show that the model's projected values and actual values agreed well, properly forecasting future water quality development trends, proving the feasibility and utility of using LSTM deep neural networks to predict drinking water quality [3].

Water quality monitoring is a crucial component of water resource management. The long short-term memory (LSTM) and convolutional neural network (CNN) models, as well as their hybrid, the CNN–LSTM model, were designed to forecast two water quality variables in the Small Prespa Lake in Greece, Dissolved Oxygen (DO; mg/L) and chlorophyll-a (Chl-a; g/L). The study's key contribution was the development of a combined CNN–LSTM model for predicting water quality characteristics. Two typical machine learning models, support vector regression (SVR) and decision tree (DT), were created to compare with the DL models. DO and Chl-a concentrations were predicted using lag durations for input variables pH, ORP, water temperature, and EC of up to one (t1) and two (t2), respectively. In both the training and testing stages, the correlation coefficient (r), root means square error (RMSE), mean absolute error (MAE), their normalised equivalents (RRMSE, RMAE; percent), percentage of bias (PBIAS), Willmott's Index, and graphical plots were employed to evaluate each model's performance (Taylor diagram, box plot and spider diagram). For DO prediction, the LSTM model beat the CNN model, but the separate DL models performed similarly for Chl-a prediction. In terms of forecasting DO and Chl-a, the standalone models were outperformed by the hybrid CNN–LSTM models (LSTM, CNN, SVR, and DT models).

By integrating the LSTM and CNN models, the hybrid model was able to capture both low and high levels of water quality indicators, particularly for DO concentrations. [4]. [5] Examines a series of AI-assisted calculations to determine the water quality index (WQI), a single record that depicts the overall water quality class (WQC) and state of water. It's an unmistakable classification based on the

WQI. pH, turbidity, total dissolved solids and Temperature are the four information boundaries used in this study. With mean absolute errors (MAE) of 1.9642 and2.7273, respectively, gradient boosting with a learning rate of 0.1 and polynomial regression with a degree of 2 were the most accurate in predicting the WQI. The multi-layer perceptron (MLP) with a configuration of (3, 7) is the most successful at classifying the WQC, with an accuracy of 0.8507. The suggested method achieves reasonable precision with a small number of parameters, increasing the likelihood of adopting real-time water quality detection.

Many ML (Machine Learning) models have become ubiquitous as they provide more accurate results than any humans. Machine learning models results can be enhanced using semantic technologies as well. Deep Learning methods have an advantage over similar problem transformation techniques on multilabel datasets as they can train on the original data without requiring data translation [6]. An ensemble comprises of a bunch of independently prepared classifiers (such as neural networks or decision trees) whose forecasts are joined while characterizing novel cases. Past research has shown that an ensemble is frequently more accurate than any of the single classifiers in the ensemble [7].

In order to minimise the harm from extreme occurrences, which would otherwise hold back development by years, weather forecasting and meteorological analysis play a crucial role in sustainable development. One of the key indications of climate change is a change in surface temperature. In this study [8], the authors propose an unique deep learning model that can accurately capture the spatial and temporal relationships of various meteorological data to anticipate temperature.

Understanding the working of complex models remains a mystery. However, these black boxes can be understood in an interpretable manner. For every black-box classifier, LIME and SHAP [9] are prominent local explanation techniques. Individual predictions for every classifier are reliably explained using these methods, surrounding each prediction, which learns an interpretable model (e.g., linear model). In particular, LIME and SHAP estimate feature attributions for each instance, which represent each feature's contribution to the black box prediction [10]. As AI fills in intricacy and effect, much expectation lays on explanation techniques as instruments to clarify significant parts of learned models. Explanations might help fulfil administrative necessities, assist experts with troubleshooting their model, and maybe, uncover bias or other accidental impacts learned by a model. [11] Refine the talk on interpretability. To begin with, they look at the inspirations' hidden interest in interpretability, discovering them to be assorted and periodically harsh. Then, at that point, they address model properties and procedures thought to give interpretability, distinguishing straightforwardness to people and post-doc clarifications as contending thoughts. All through, they examine the feasibility and desirability of various ideas and question the frequently made declarations that linear models are interpretable and that deep neural networks are not.

[12] Presents the details of existing procedures for ensemble classifiers, and shortcomings of various understanding methodologies. They additionally talk about pivotal issues that the classification models will have to consider in future work, for example, planning easy-to-use clarifications and creating extensive assessment measurements to further push forward the space of interpretable AI. For handling unbalanced dataset problems, a two-step supervised learning approach based on a single layer feed forward Artificial Neural Network (ANN) is proposed in this paper [13].

3. Proposed Work

In this section the dataset processing and proposed models are discussed in detail. Six features have been considered that help for classifying the water quality based on CCME WQI.

3.1. Data Pre-Processing

Data pre-processing is an integral part of machine learning, data mining or any task related to data science. Data pre-processing is a technique that can be used to change the raw data i.e unclean data gathered from various sources to clean data that can be used for analysis, data mining, machine learning, etc. The data that are received may not be free from redundancy, outliers, or null values but there can be many extreme values present in the data which can create anomalies in the predictions. Data pre-processing has been divided into 4 stages:

1) Data cleaning: filling null values, smoothing noisy data, removing inconsistencies in the data.
2) Data integration: adding different types of data to facilitate uniformity of data using multiple databases from different type of data
3) Data reduction: removing the outliers or corner cases to improve the accuracy of the model
4) Data transformation: transforming data so that it can be used by various machine learning models. Sometimes the data present may be imbalanced i.e., some classes have far lesser values as compared to some other classes. In that case, the accuracy may be higher but prediction of minority classes may suffer as the data of minority class is less. There are various ways to tackle this problem. Some of them are random under sampling, random over sampling, edited nearest neighbour and SMOTE. The dataset have been divided into 5 different classes according to CCME WQI as shown in Table 1.

TABLE 1: RANGE OF WQI

Quality	Excellent	Good	Fair	Marginal	Poor
WQI Range	95-100	80-94	65-79	45-64	0-44
Class	1	2	3	4	5

In an ideal situation, the classes are balanced, implying that the number of instances for each class is nearly equivalent. However, some real-world databases lack this property, resulting in unbalanced classes [14-16]. In this dataset some classes are having less data values as compared to other classes. So, to tackle this problem, SMOTE is used. SMOTE stands for Synthetic Minority Oversampling Technique. Oversampling refers to copying or creating new examples of the minority classes in the dataset so that the number of examples in the minority class is close to the number of examples in the majority classes so that the prediction accuracy can be increased. SMOTE starts by looking for examples that are similar in terms of features. It then draws a line in the feature space between the instances and a new sample point along the line. First, a representative from the minority class is chosen at random. Then, given the example we've picked, k parameters of the nearest neighbour are discovered (usually k=5). A randomly chosen neighbour is then chosen from the examples, and a new example is repeated at a randomly determined location in the feature space between the two instances it selected.

3.2. Artificial Neural Network

An artificial neural network (ANN) is a type of artificial intelligence that mimics the functions of the human brain. An artificial neuron is designed to work like a biological neuron by performing operations on the values of the input it receives. If the values are above some threshold, the artificial neuron sends its own signal to outputs, which is then received by other artificial neurons. In a neural network each neuron receives input, analyses it using a specific function and then transmits the result further. The way a network is built is called network construction. The network is constructed with various layers that provide enough learning. The input layer is the first of three layers. The output layer is the last one. The hidden layer is the entire layer in the middle, and there can be as many hidden layers as required. In this network each neural node connects to the rest of the node in the next layer. Each route from one node to another carries weight which forms the basis of propagation of input. One would think that for a signal to pass from one place to another it would have to perform the mathematical calculation of multiplying its value in the initial phase by the weight mentioned of the path and this product would eventually reach the end of the node. Each node in the network receives several of these products from the nodes in the previous layer they are connected to. Integrates them all, executes the function and transfers the output to the next layer where the process repeats. Neural networks learn by adjusting the weight values of each method so that the output layer produces an impression that can mimic your data. In this way, complex incompatible hypotheses can be created and the most hidden layers with the most powerful models can be created. ANN is used in this study to recommend and anticipate future trends in selected higher education computer science/technology courses [17]. Parameters for the artificial neural network:

1) Learning rate: This hyper parameter regulates how much the model can change in response to the estimated error each time the model ratings are altered.
2) Batch size: The amount of training examples used in one iteration is referred to as batch size. The entire dataset is separated into a number of equal batches.
3) A train step means one gradient update. In one step, one batch of data is processed. The no. of training steps is equivalent to the iteration.
4) Epoch: One epoch is said to complete, when the whole input is processed through the neural network.
5) The role of Optimizers is to use techniques or procedures for reducing losses by changing aspects of your neural network, such as weight and learning level.
6) Adam optimizer: Adam (Adaptive Moment Estimation) is a first-order and second-order strength optimizer. Adam's thought is not to roll over too rapidly in-depth, but to slow down a little to search more carefully.
7) Activation function: it examines the Y value produced by the neuron and determines whether external communication should view this neuron as "fired" or not. Relu function is used in this work.

3.3. XG-Boost

Tree boosting is a profoundly powerful and broadly utilized AI technique. It is used by many researchers to achieve state-of-art results on many machines learning challenges Boosting trains various powerless classifiers consecutively on contrastingly weighted variants of training samples, while, bagging trains various classifiers autonomously on bootstrap samples. XGBoost (Extreme Gradient Boosting) is a gradient boosting based on ensemble Machine Learning technique. It is highly flexible, efficient and portable neural networks tend to beat any algorithm in case of prediction problems involving unstructured data like text, images and so forth. Nonetheless, with regards to small to medium structured /tabular data, decision tree-based algorithms are viewed as top tier at this moment [18][19][20].

Since a medium scale tabular dataset is used for classification, few options based on ensemble learning like bagging, boosting, etc can only be opted. When the goal is to reduce the variance of a decision tree classifier, bagging is utilised. The goal of bagging is to construct several subsets of data from training samples selected at random with replacement. Their decision trees are trained with each group of data. The classification model will not have accurate values because the final prediction is based on the mean predictions from subgroup trees. A set of predictors are established as number increases. In this approach evaluation is based on feedback from previous predictors. It boosts efficiency. So, XGBoost have been chosen for classification as it is more suitable for the use case. Parallel processing, handling missing values, Regularization, tree pruning, high flexibility, and built-in cross validation are just a few of the benefits.

3.4. Explainable AI

Most machine learning algorithms are black-box models that mean how the algorithm or neural network selects or denies some features are not known. As more and more innovations are taking place in the field of neural networks, this needs to be figured out exactly how it works. Explainable AI works towards that goal, and it provides explanations to different ML models by facilitating a global understanding for humans. This branch of AI is an emerging field in machine learning that helps to address divergent black-box models [21].

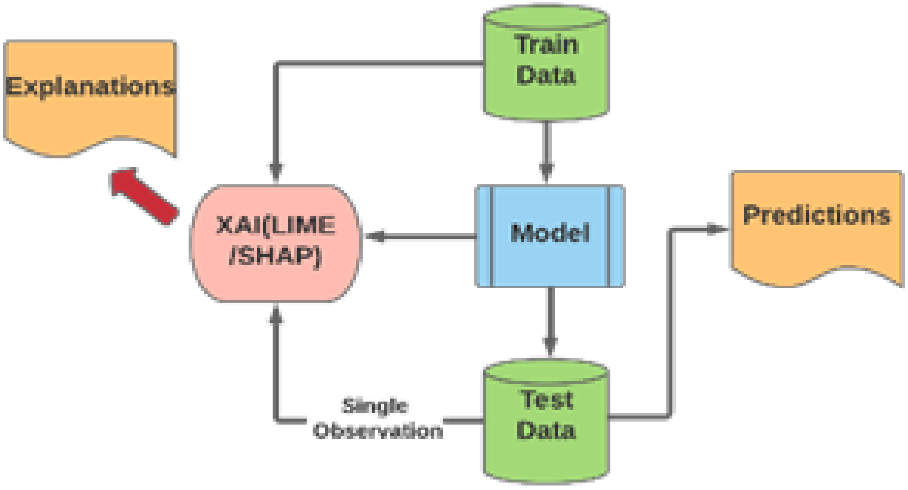


Figure 1: Work Flow

The normal workflow of the machine learning model is where the learning takes place on the training data and maps the inputs with the desired target variable. Then the model makes predictions for the test data. In order to understand complex black-box models, XAI (explainable AI) tools like LIME & SHAP utilize the training dataset alongside the model to clarify a specific forecast made on the test data as shown in Figure 1.

LIME stands for Local Interpretable Model-Agnostic Explanations. Model agnosticism means that it can explain the predictions for any supervised model in an interpretable manner. Local explanations infer that these explanations given are constant within the locality of the sample provided. Feature importance tells which features are predominant on a dataset level but is difficult to diagnose specific prediction models. Lime provides local model interpretability that addresses questions like; which feature value affected the prediction value? why did the model give this prediction? The behaviour of the model is explained in that neighbourhood rather than as a whole. LIME tries to figure out what makes a classification model work by changing the input variables and seeing how the predictions change [21]. It provides explanations that are innate to explain to users and works for both structured and unstructured data. Randomly select a single data point and generate a dataset of perturbed points. Then the predictions for each data point are derived that give some idea about the decision surface. The number of features is selected to provide a suitable explanation. Then the explanation model is computed using the predictions.

SHAP (Shapley Additive exPlanations), an XAI tool derived from the Shapley values. SHAP helps us to predict how each feature has contributed to the model prediction. The average marginal contribution of a feature value across all conceivable permutations is the Shapley value. The advantages of using SHAP are:

1) Global Interpretability - the positive or negative relationship of each feature with the target variable is shown by the SHAP values.
2) Local interpretability - explains each individual prediction where each observation gets a set of SHAP values.
3) Usability across any tree-based model which performs fast computations.

4. Results and discussion

After data pre-processing and making use of SMOTE for increasing minority class distribution and balancing the data, parameters (features) were passed the as inputs for the ANN as well as for the XGBoost model. Hyper parameters like multi-SoftMax have been fine-tuned, implementing multiclass classification with max-depth as 4 (it should be between 3-10) since more depth value might over fit the model. N-estimators of 2000 have been chosen and "num_classes" as 5. After fine-tuning, test accuracy around 90% is got. After tuning the hyper parameters of the ANN model with 3 dense layers, relu as activation function and added drop out to the ANN model. Adam optimizer is used and call backs also used to save the model to have least loss. Figure 2 and 3 shows the performance of the ANN model and table 2 shows different metrics for the aforementioned models to indicate their performance on the dataset.

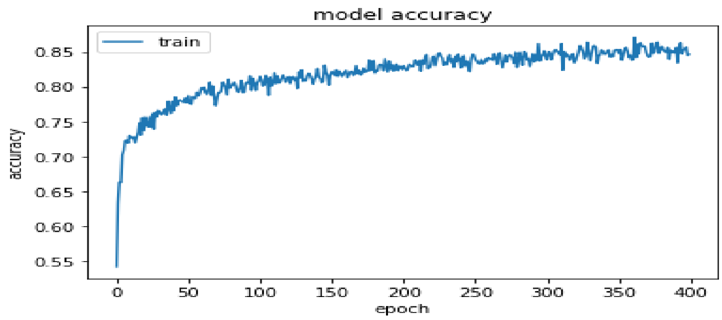


Figure 2: Training accuracy of ANN model

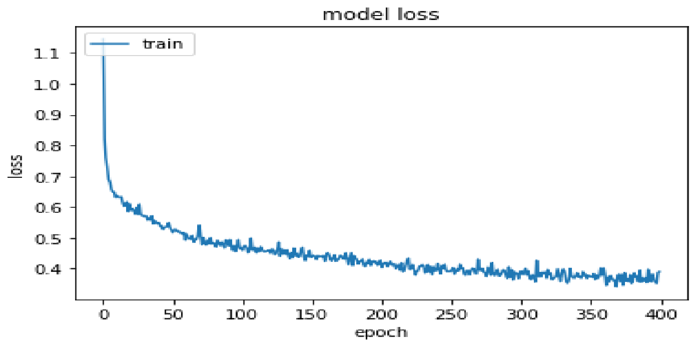


Figure 3: Training Loss of ANN

We found that ANN and XGBoost performed better than other algorithms [22] for predicting water quality with an accuracy of 0.8614 and 0.8987 as shown in Table 2.

Table 2

Classification Results Comparison

Dataset	Algorithm	Accuracy
PCRWR (Pakistan Council of Research in Water Resources)	KNN	0.7270
	Random Forest	0.7587
	Guassian Naïve Bayes	0.7843
	SVM	0.7979
	Gradient-Boosting Classifier	0.8130
	MLP	0.8507
WQI Parameter Scores 1994-2013	ANN	0.8614
	XGBoost	0.8987

4.1. Lime

Machine learning models results can be enhanced using semantic technologies [23] as well. TabularExplainer have been considered for the tabular data, which is a combination of columns. The lime.lime_tabular has various parameters such as training_data, feature_names, training_labels, class_names and the mode is specified for this instance. The explain_instance is used for explaining by inputting the instance, the predicted method of the trained model, top labels and the number of features needed for the description. The target values are different water-quality index and chosen are ph, TPN, FC, Temp, TSS & Oxy as the feature variables for water quality classification.

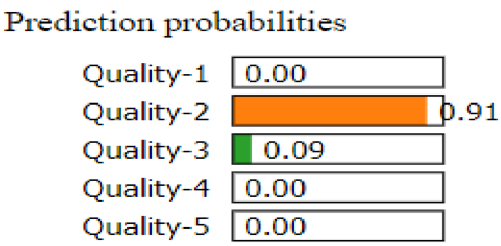
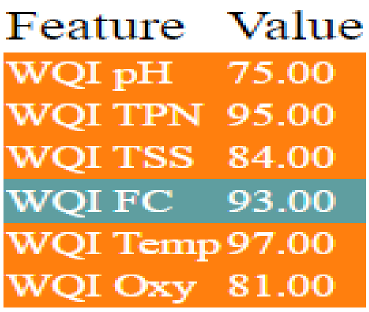


Figure 4: Prediction probability for each class



Feature	Value
WQI pH	75.00
WQI TPN	95.00
WQI TSS	84.00
WQI FC	93.00
WQI Temp	97.00
WQI Oxy	81.00

Figure 5: Actual values of different features

Figure 4, shows the prediction probabilities for each class and Figure 5 contains the actual values of the six features as color-coded ones. Figure 6, 7 and 8 gives explanation of features for quality values 1, 2 and 3 respectively.

Figure 6: Explanation of features for Quality-2

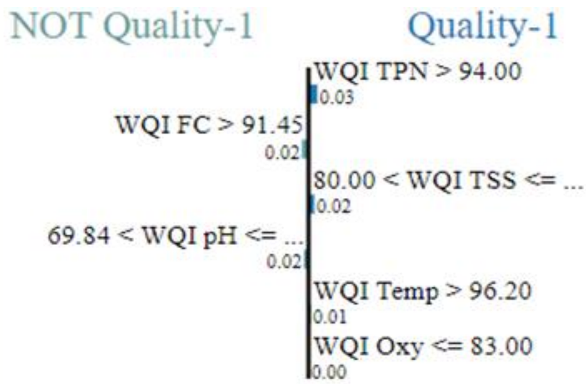


Figure 7: Explanation of features for Quality-1

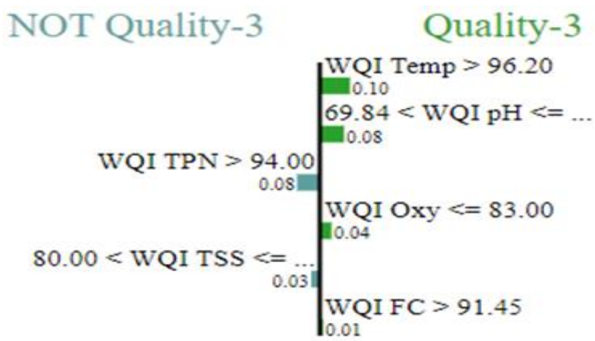


Figure 8: Explanation of features for Quality-3

In Figure 5, the features such as the WQI pH, WQI TPN (nitrogen), WQI Temp (temperature), WQI Oxy (oxygen), and WQI TSS (total suspended sediment) constitutes ninety-one percentages for water quality-2 index. The feature values highlighted in orange contribute to the above water quality index. These values fall within the definite range that helps for this classification. WQI FC (fecal coliform bacteria) has a value greater than 91.45 that negatively contributes to water quality-2. There is a nine percent probability that this instance can be grouped under quality-3 highlighted in green color. The features that contribute to quality-3 are WQI Temp, WQI pH, and WQI oxy.

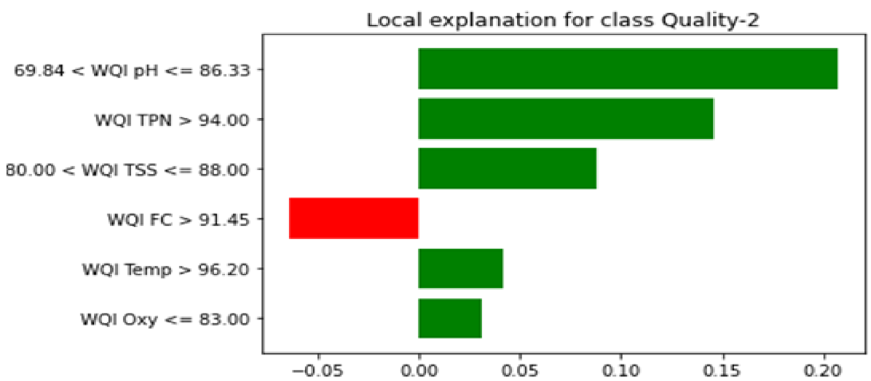


Figure 9: Local Explanation for class Quality-2

The exp.as_pyplot_figure() plots the correlations of each feature value with the target. Figure 9 depicts the positive correlations of these feature variables with the target are depicted in green, otherwise in red. WQI TPN> 94: high nitrogen values positively correlate with high water quality. WQI FC>91.45: high fecal coliform bacteria values negatively correlate with high water quality. WQI Oxy<=83.00: low oxygen values positively correlate with high wine quality.

4.2. SHAP

Two different models are used in the work, for water quality classification. For ANN model, shap.DeepExplainer have been used which is an upgraded version of the DeepLIFT algorithm [21]. This explainer approximates SHAP values for deep learning models and takes parameters such as model and data. Then the SHAP values are calculated for the observation data.

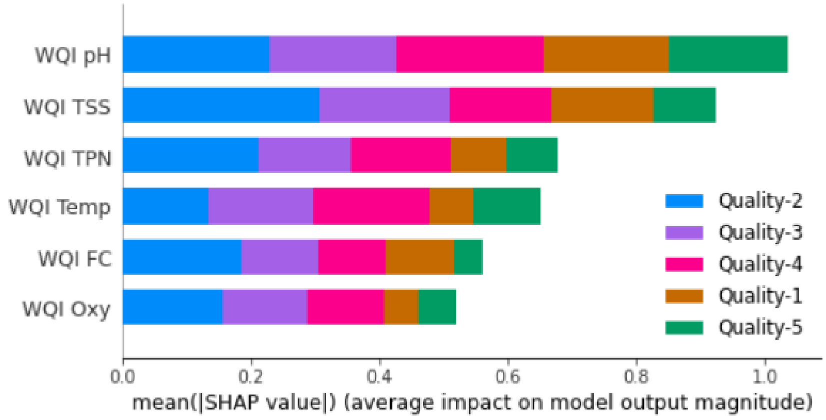


Figure 10: SHAP Summary Plot for ANN

The shap.summary_plot provides an insight about the model to know about feature importance. The x-axis represents the mean absolute Shapley values of each instance and the y-axis lists the feature variables. In figure 10, WQI pH is the crucial feature that has a high Shapley value range and others are listed based on their importance. This has a high impact for the prediction of the target variable. Class 4 and 5 hardly uses the WQI FC feature and other classes use the other features equally.

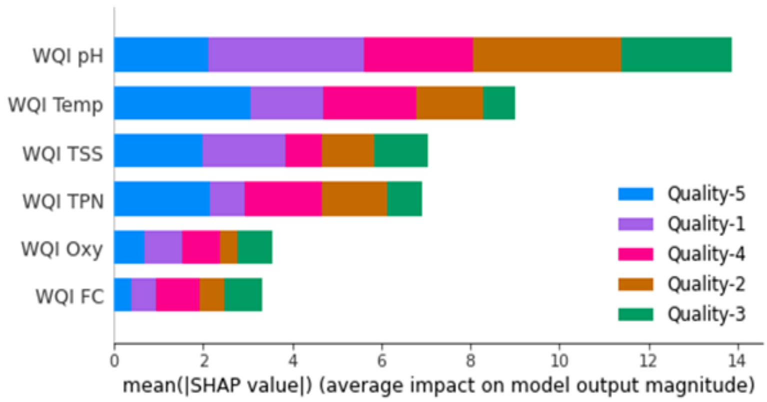


Figure 11: SHAP Summary Plot for XGBoost

In figure 11, quality 2 and quality 5 hardly use the WQI FC feature and other classes use the other features equally.

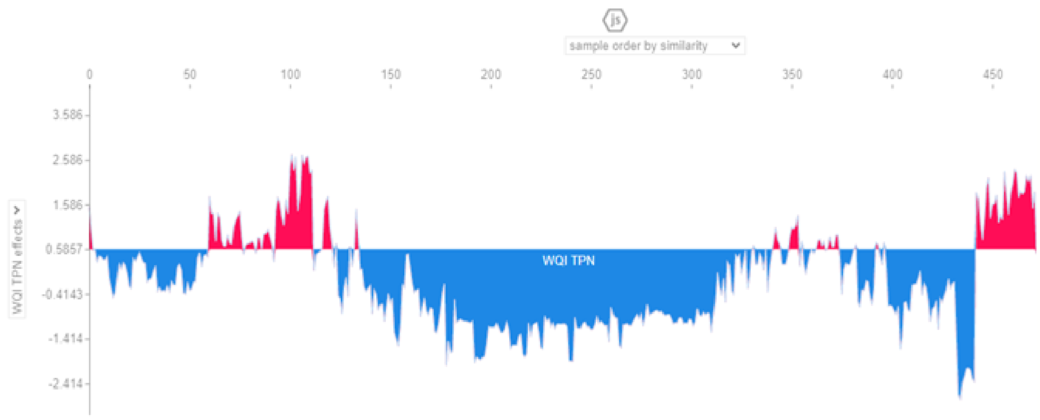


Figure 12: SHAP Force Plot

SHAP force plot is a stacked SHAP explanation clustered by explanation similarity as shown in Figure 12. Each position on the x-axis represents an instance of the data (value of the metric). Red SHAP values in Figure 12 represent the increase in the prediction, blue SHAP values decrease it. On the right side of Figure 12, there is a cluster that represents a high prediction for quality 2. Table 3 lists the differences between Lime and SHAP in terms of their usage.

Table 3 Lime and SHAP

Feature	Lime	SHAP
One-hot Encoding	For a dataset that is already one-hot encoded, it may respond such that the same feature has two values.	Encoded data are interpreted properly.
XgBoost	Can't deal with XGBoost algorithm to utilize xgb.DMatrix() on the information for classification.	SHAP follows a progressive strategy to deduce results from tree-based models.
Computation Time	LIME is fast, however inconsistent explanations are observed rarely.	SHAP is slow, but behaves better in giving consistent explanations
Probabilistic classification	LIME works with models that yield probabilities for classification issues. This might bring some bias into the explanations.	KernelSHAP provides explanations in general by collecting information across many instances.
Uses	Lime works better for single prediction explanation.	SHAP is suitable for entire machine learning model explanation

5. Conclusion

Explainable AI is a field of AI where understanding the results of the solution is made comprehensible to humans. Most machine learning algorithms are "black-box" in which the programmer himself does not understand how his model arrived at a particular solution or a specific decision. In this work, Explainable AI usage on real-world applications is discussed. According to our ANN & XGBoost algorithm, we found that ph & temp respectively were the most significant factors that contributed to bad water quality (quality 5). The increased temperature may not be tolerable for aquatic life as it increases microbial growth, which in turn decreases dissolved oxygen, makes metals more bioavailable, or in other ways increases the harm from nutrients and toxins. When pH levels fall outside of this range (up or down), animal systems are stressed, and hatching and survival rates suffer. We have used the water quality dataset to demonstrate how XAI (explainable AI) works and know how various features contribute to a particular decision arrived at by the machine learning models, in turn, can prevent water pollution that is an immense problem in the current era.

References

[1] PCRWR. National Water Quality Monitoring Programme, Fifth Monitoring Report (2005–2006); *Pakistan Council of Research in Water Resources Islamabad*: Islamabad, Pakistan, 2007.

[2] A. Solanki, H. Agrawal, and K. Khare, "Predictive analysis of water quality parameters using deep learning," *International Journal of Computers and Applications*, vol. 125, no. 9, pp. 29–34, 2015.

[3] Liu, P.; Wang, J.; Sangaiah, A.K.; Xie, Y.; Yin, X, "Analysis and Prediction of Water Quality Using LSTM Deep Neural Networks in IoT Environment" , *Sustainability* 2019, 11, 2058.

[4] Barzegar, R., Aalami, M.T. & Adamowski, J. "Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model", *Stoch Environ Res Risk Assess* 34, 415–433 (2020).

[5] Ahmed, U., Mumtaz, R., Anwar, H., Shah, A.A., Irfan, R., & García-Nieto, J. (2019). "Efficient Water Quality Prediction Using Supervised Machine Learning", *Water*, *11*, 2210.

[6] Maxwell, Andrew, Runzhi Li, Bei Yang, Heng Weng, A. Ou, H. Hong, Zhaoxian Zhou, P. Gong and C. Zhang. "Deep learning architectures for multi-label classification of intelligent health risk prediction." *BMC Bioinformatics* 18 (2017): n. pag.

[7] David Opitz, Richard Maclin, "Popular Ensemble Methods: An Empirical Study", *Journal of Artificial Intelligence Research,* (1999) 169-198

[8] M. A. R. Suleman and S. Shridevi, "Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model," in IEEE Access, vol. 10, pp. 82456-82468, 2022, doi: 10.1109/ACCESS.2022.3196381.

[9] Scott M Lundberg and Su-In Lee. 2017,"A Unified Approach to Interpreting Model Predictions", *In Neural Information Processing Systems (NIPS),* Curran Associates, Inc., 4765–4774.

[10] Dylan Slack, Sophie Hilgard, Emily Jia, Sameer Singh, and Himabindu Lakkaraju, "Fooling LIME and SHAP: Adversarial Attacks on Post hoc Explanation Methods" *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 180–186, 2020.

[11] Zachary C. Lipton, "The Mythos of Model Interpretability", *Workshop on Human Interpretability in Machine Learning* (WHI 2016), New York, NY

[12] Jana A, Krishnakumar SS. Sign Language Gesture Recognition with Convolutional-Type Features on Ensemble Classifiers and Hybrid Artificial Neural Network. Applied Sciences. 2022; 12(14):7303. https://doi.org/10.3390/app12147303

[13] Adam, Asrul & Shapiai, Mohd Ibrahim & Chew, Lim & Ibrahim, Zuwairie & Jau, Lee & Khalid, Marzuki & Watada, Junzo. (2010)," A Two-Step Supervised Learning Artificial Neural Network for Imbalanced Dataset Problems", *International Journal of Innovative Computing, Information and Control.* (IJICIC).

[14] Estabrooks, A., Jo, T., Japkowicz, N., "A Multiple Resampling Method for Learning from Imbalanced Data Sets", *Computational Intelligence* 20, 18–36 (2004)

[15] Chawla, N.V., Japkowicz, N., Kotcz, "A.: Editorial: special issue on learning from imbalanced data sets", *SIGKDD Explor. Newsl*. 6, 1–6 (2004)

[16] Sun, Y.M., Wong, A.K.C., Kamel, M.S.: "Classification of imbalance data: A review", *International Journal of Pattern Recognition and Artificial Intelligence* 4, 687–719 (2009)

[17] Tianqi chen, carlos guestrin,"XGBoost: A Scalable Tree Boosting System", *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* , August 2016 Pages 785–794

[18] Chowdhury, Dilip & Sen, Deepanjan. (2017). "Artificial Neural Network Based Trend Analysis and Forecasting Model for Course Selection", *International journal of computer sciences and engineering.* 05. 20-26. 10.5281/zenodo.5226838.

[19] P. Saran, D. Rajesh, H. Pamnani, S. Kumar, T. G. Hemant Sai and S. Shridevi, "A Survey on Health Care facilities by Cloud Computing," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-5, doi: 10.1109/ic-ETITE47903.2020.231.

[20] Md AQ, Jaiswal D, Daftari J, Haneef S, Iwendi C, Jain SK. Efficient Dynamic Phishing Safeguard System Using Neural Boost Phishing Protection. Electronics. 2022; 11(19):3133. https://doi.org/10.3390/electronics11193133

[21] Ullah, hsan & Rios, Andre & Gala, Vaibhav & Mckeever, Susan. (2020),"Explaining Deep Learning Models for Structured Data using Layer-Wise Relevance Propagation"

[22] Ahmed, Umair & Mumtaz, Rafia & Anwar, Hirra & Shah, Asad & Irfan, Rabia & García-Nieto, José. (2019). Efficient Water Quality Prediction Using Supervised Machine Learning. Water. 11. 2210. 10.3390/w11112210.

[23] Rachana L., Shridevi S. (2021) A Literature Survey: Semantic Technology Approach in Machine Learning. In: Zhou N., Hemamalini S. (eds) Advances in Smart Grid Technology. Lecture Notes in Electrical Engineering, vol 688. Springer, Singapore. https://doi.org/10.1007/978-981-15-7241-8_34 , First Online 19 September 2020.