ISBN: 978-628-96613-1-6. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.752

Stacking ensemble model with heterogeneous algorithms for the prediction of the water quality index of the Rimac basin

Mg. Jose Briones- Zuñiga[1]; Dr. Juan Soria Quijaite[2]
[1,2]Universidad Tecnológica del Perú, Perú, *C19980@utp.edu.pe, C20723@utp.edu.pe*

Abstract*– Water quality monitoring is essential for the protection of public health and ecosystems. This research used historical data of the physicochemical and microbiological parameters of the Rimac River basin in the city of Lima, Peru, from 2014 to 2021, and proposed a stacking ensemble model with heterogeneous algorithms for the prediction of the water quality index (NSF) in the Rimac River basin/Peru. The results show low values of the mean square error (MSE) and mean absolute error (MAE) of 9.954 and 2.433 respectively. Likewise, a high level of fit with a coefficient of determination of 85.9%. The selection of the prediction model algorithms was based on the detection of stationarity and autocorrelation in the target variable - water quality index. It is concluded that it is necessary to strengthen and use the heterogeneous algorithm to predict the water quality of the Rimac basin. It was developed in a Google Colab environment and Python programming language

***Keywords*-- Stacking assembly; water quality index; algorithms; prediction model; machine learning.**

I. INTRODUCTION

The waters on the surface of the planet are sensitive to contamination, whether natural or by human activity over time. Therefore, it is essential to investigate the evaluation of water quality in order to take actions to control and reduce the deterioration of water supply sources, with the objective of providing safe water [1]. In addition, high levels of pollutants such as nitrogen and phosphorus can cause eutrophication, which affects water quality and causes gastrointestinal and respiratory diseases [2]. the presence of harmful bacteria and viruses in contaminated river water can lead to outbreaks of waterborne diseases. In an investigation in the Yamuna River, high levels of fecal coliforms have been associated with unexpected outbreaks of bacteria that create health risks [3]. Also, rivers are essential for irrigation and industrial processes. It is not appropriate to use contaminated water for these purposes, which reduces the available water supply and affects food production and industrial operations [4]. Various pure and hybrid water quality predictive models of different nature have been experimented with as a monitoring and prevention tool. The models often tend to be tailored to specific river basins or regions, sensitive to local environmental factors and pollution sources. In recent research, models developed for the Bhavani River in India and the Minjiang River in China incorporate local water quality indicators and pollution sources [5][6].

Research in China, Kerala and Tamil Nadu has shown that the application of machine learning models, such as temporal graph convolutional neural network (T-GCN) and support vector machines (SVM), achieve high accuracy in predicting water quality data from multiple monitoring stations [7][8]. Likewise, spatial stream network (SSN) models, which use hydrological distance and topological data structures, emerge as novel techniques for predicting water quality [9]. On the other hand, research in India has used artificial neural networks (ANN), multivariate adaptive regression spline (MARS) and least squares support vector machine (LS-SVM) as machine learning methods for predicting water quality parameters with good results [10].

In a research in Atlanta, USA, it was shown that hybrid process-based watershed models and an artificial neural network (ANN) generated optimal results for the prediction of water quality parameters in unmonitored watersheds [11]. On the other hand, in a research they used machine learning models such as multivariate adaptive spline regression (MARS), least squares support vector machine (LS-SVM) and decision trees (DT) with the objective of predicting water quality indices generated results with high accuracy [12]. Similarly, an investigation demonstrated that the ANN model is suitable for monitoring water quality conditions, as its results showed small deviations for several water quality parameters of less than 5% [15].

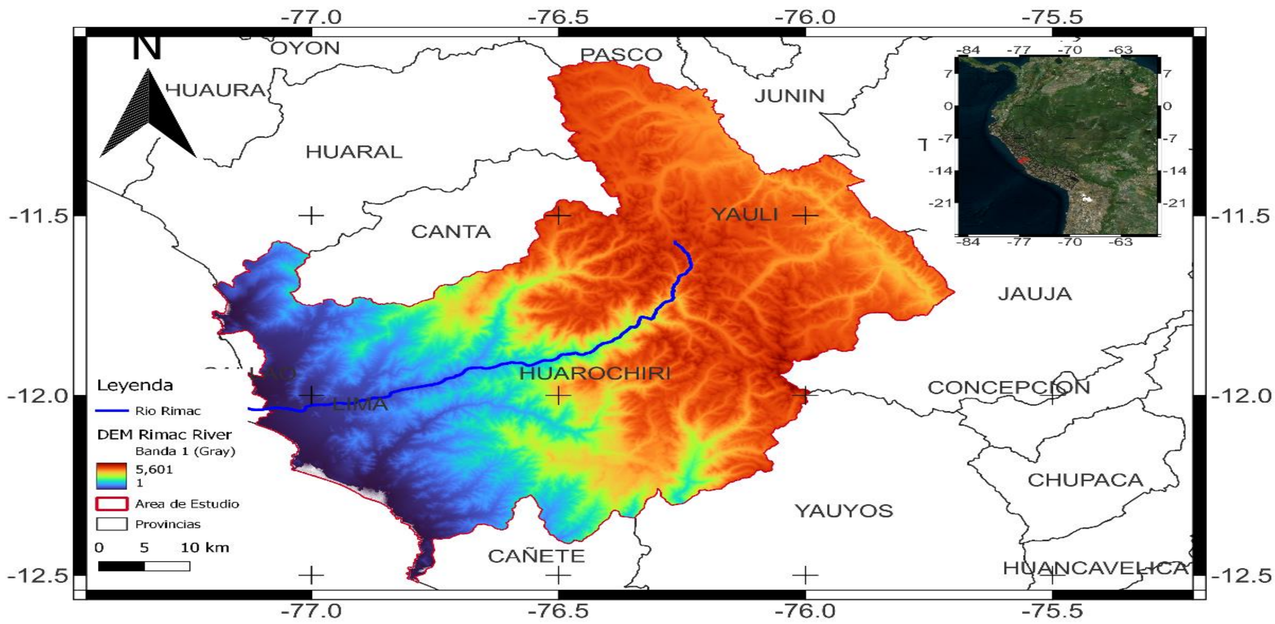

Fig. 1 Location map of the Rimac basin

Likewise, research has been carried out using different types of supervised learning algorithms such as neural networks (ANN), support vector machines (SVM), decision trees (DT), Naive Bayes (NB) and K nearest neighbors (KNN), linear and nonlinear regression models, as well as Monte Carlo simulations or hydrological simulation programs such as FORTRAN with the objective of modeling water quality for monitoring and reducing uncertainty[16][17][18].

Likewise, in various investigations to evaluate the performance of prediction models, prediction error metrics such as mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) and as an indicator of goodness of fit the coefficient of determination (R^2) of river water quality prediction models have been used [7] [11] [12] [13]. On the other hand, in an investigation in Mexico, the PBIAS index was used to evaluate the performance of the water quality prediction model, since it incorporates the uncertainty inherent in the water quality data [14].
In previous studies, the different criteria for supervised learning algorithms and traditional models used for predictive purposes of water quality were evidenced, but the absence of algorithm selection based on the characteristics and nature of physicochemical and microbiological data as a result of pre-modeling analysis was identified.

The objective of the research is to develop a stacking ensemble model with heterogeneous algorithms to predict the water quality index using the NSF-WQI methodology. This approach is based on the combination of multiple base models of a bagging nature and neural networks and a boosting metamodel to generate more robust and accurate predictions, allowing a better assessment of water quality based on the physicochemical parameters of the Rímac basin.

II. METHODOLOGY

A. *Study site*

The study site uses information from the Rimac basin with jurisdiction Chillon – Rimac – Lurin located in the city of Lima/Peru, as shown in Figure 1.

B. *Data selection*

The data set used for the development of the ensemble stacking prediction model was constructed from data provided by the National Water Authority (ANA). These data include the physicochemical and microbiological water parameters necessary for the calculation of the NSF-WQI water quality index. A weighting was applied to each parameter, as shown in Table I, to determine its relative importance. Equation 1 was used to calculate the target variable (Final WQI) of the water quality index, which allowed the generation of a complete and accurate data set [1].

TABLE I

PARAMETER WEIGHTING (NSF-WQI)

.

Parameter	Weighting
Nitrates	0.10
pH	0.12
Turbidity	0.08
Phosphates	0.10
Dissolved oxygen	0.17
DBO5	0.10
Fecal coliforms	0.15
Temperature	0.10

The average NSF-WQI water quality index is calculated using (1)

$$\text{Final QI} = \sum_{i=1}^{n} w_i * q_i \qquad (1)$$

Where w_i is the weighting of parameter i and q_i is the quality of the parameter i.

The NSF-WQI formula calculates a numerical value that represents the overall water quality. Each water quality parameter q_i is weighted by its relative importance w_i. The sum of the products of the weights and quality values provides a measure of water quality. This index is useful for comparing water quality at different locations or times, facilitating the identification of changes or trends in water quality [1].

The first step consisted of obtaining data from the reports of the National Water Authority (ANA) corresponding to the Rimac basin and the Chillon-Rimac-Lurin jurisdiction, covering the eight-year period between 2014 and 2021. These reports provided the necessary data on physicochemical and microbiological parameters that were used as inputs for the model; this range of years was delimited due to the lack of accessibility of data from recent years by the institution in charge of monitoring. According to, it is shown in Table I. The data were manually extracted from tables and reports presented in ANA documents, ensuring that they covered different monitoring stations throughout the watershed, to obtain an adequate representation of water quality.

In the second step, a data consistency process was carried out, which included a statistical analysis. During this process, missing values and outliers are identified through statistical analysis and data visualization. Methods such as outlier detection and the use of boxplots were employed to identify outliers.

The third step was the creation of the target variable WQI (Water Quality Index), which was calculated using the expression (1) that combines the values of the exogenous variables. This index provides an overall measure of water quality, integrating all parameters into a single metric. The calculation of the WQI was performed following the standard methodology that assigns different weights to each parameter according to its relative importance in water quality. It was also implemented using the Python programming language.

The entire process of data matrix construction and WQI calculation was carried out on the Google Colab platform, taking advantage of Python libraries such as pandas for data manipulation and cleaning, numpy for mathematical operations, and matplotlib for data visualization.

C. Data preprocessing

Min-Max normalization is an essential technique in data preprocessing, particularly useful to standardize the values of physicochemical and microbiological parameters in a specific range, between 0 and 1. This transformation ensures that all features have the same scale, which is crucial for supervised learning models such as neural networks and allows for improved convergence and training. As shown in (2)

.

$$x' = \frac{x_i - min\ (x)}{max\ (x\ \ ) - min\ (x)} \qquad (2)$$

Where x_i represents each individual value of the characteristic x, min(x) is the minimum value of x y max(x) is the maximum value of x. This mechanism facilitates interpretation and allows comparison of data, maintaining the integrity of the original information in a standardized and consistent range.

D. Stacking assembly model with heterogeneous algorithms

The selected strategy for the prediction of the water quality index is the ensemble stacking method with heterogeneous algorithms, an advanced machine learning technique that combines bagging, boosting and neural network models in order to improve the accuracy and robustness of water quality index (Final WQI) predictions based on physicochemical and microbiological parameters. The prediction method is divided into two levels:

- *Base Models:* The Decision Tree, Random Forest and multilayer perceptron base models are trained to generate individual predictions for the final WQI target variable using the input data to the physicochemical and microbiological parameters.

- *Metamodel:* The predictions from the base models are used as the new dataset for the metamodel (XGBoost) in order to generate a more accurate final WQI prediction. This mechanism enhances the ability to capture the strengths of each base model and reduce their individual weaknesses, resulting in a more robust and accurate model [19].

Decision Tree

Decision Tree is a supervised learning algorithm of bagging nature that partitions the feature space into smaller sectors based on decision rules. Each internal node of the tree represents a condition on a feature, and the leaves reflect the predictions (in this case, the Final WQI variable). On the other hand, a disadvantage is that they tend to overfit if their depth is not controlled [20].

Random Forest

Random Forest is a model based on decision trees that combines multiple trees trained using bootstrapped sampling of the data and features (random feature selection). The final prediction is the average of the predictions of the individual trees in the specific case of regression. Random Forest reduces the likelihood of overfitting and handles non-linear and high-dimensional data well [21].

Multilayer Perceptron (MLP)

The architecture of the multilayer perceptron (MLP) neural network is an interdependent structure that contains layers of neurons energized by internal algorithms that optimize the parameters to minimize their error margin, inspired by the biological neural network. Each neuron uses a non-linear activation function on the linear combination of the product of its inputs by its synaptic weights. The MLP has the ability to model complex and non-linear relationships between the explanatory variables and the target variable (Final WQI). However, it requires an experimentation process with the aim of adjusting the hyperparameters [22].

XGBoost

The XGBoost algorithm (eXtreme Gradient Boosting) is a boosting algorithm based on decision trees that optimizes a loss function using gradients. Likewise, XGBoost builds trees sequentially, where each new tree reduces the residual errors of the previous ones. This makes it an ideal algorithm for predicting the target variable, especially in structured or tabular datasets. XGBoost is known for its efficiency, scalability, and ability to handle missing data[23].

E. Prediction model performance evaluation metrics

In order to measure the performance of the water quality index prediction stacking ensemble model, three essential metrics were used: MSE (Mean Squared Error), MAE (Mean Absolute Error) and Coefficient of Determination (R^2) [24]. The MSE measures the magnitude of the squared errors, penalizing larger errors. As shown in (3)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (3)$$

The MAE calculates the average of the absolute errors. As shown in (4)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (4)$$

The R^2 measures the ability of the model to capture the variability of the data; a value close to 1 indicates a good fit[11-13]. As shown in (5)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2} \quad (5)$$

Fig. 2 presents the diagram that sequentially describes the stages of the stacking ensemble method with heterogeneous algorithms estimated to predict the water quality index.

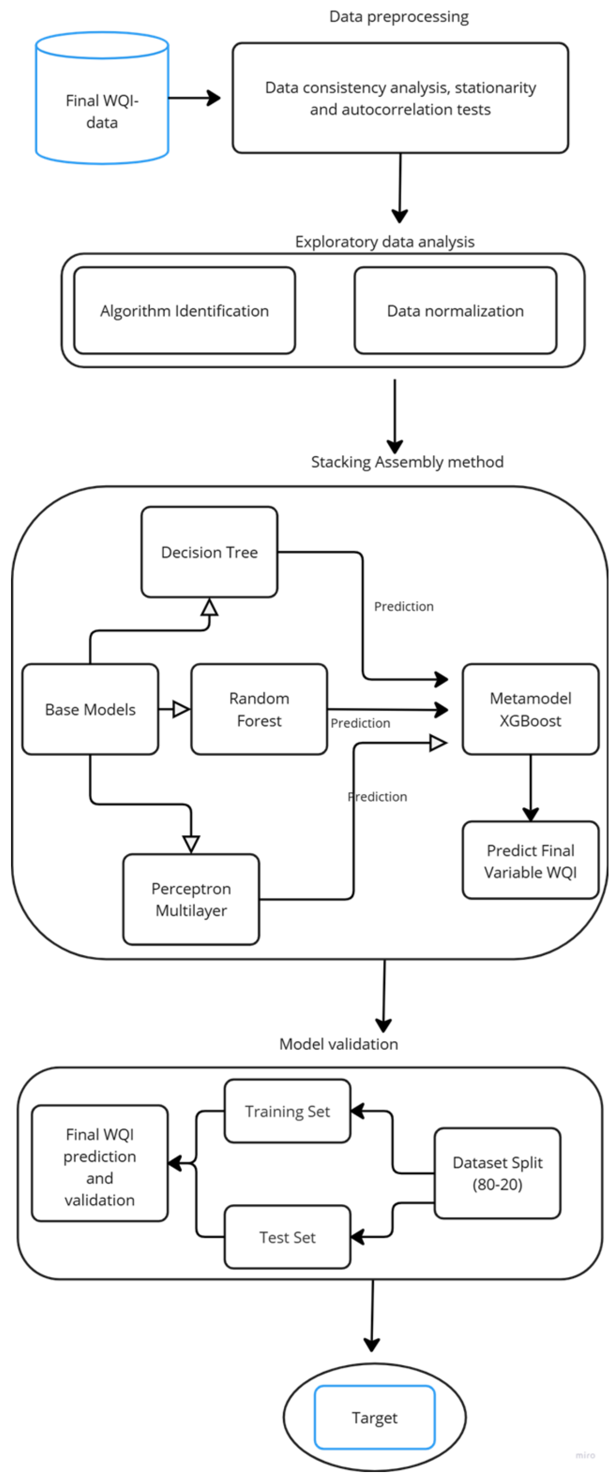

Fig. 2 Flowchart of the ensemble stacking method with heterogeneous algorithms for water quality index prediction.

III. Results and Discussion

A. Descriptive statistics on physicochemical and microbiological parameters

Table II shows the main statistics of the physicochemical and microbiological parameters for calculating the water quality index of the Rimac basin. There are indications of high levels of contamination due to the high variability of the biochemical oxygen demand (BOD) and Ecoli. Likewise, the averages of these parameters are above the limits, evidence of organic and bacterial contamination. On the other hand, pH values are close to the upper limit.

TABLE II

STATISTICS OF PHYSICOCHEMICAL AND MICROBIOLOGICAL PARAMETERS

Statistic	pH	BOD mgL	DO mgL	Fosfato	Ecoli	Nitrato
Mean	8.00	26.8	6.7	0.23	75,431.89	1.83
Std	0.81	49.2	1.5	0.66	477,476.65	1.55
Median	8.22	8.6	7.2	0.02	170.00	1.57

B. Analysis of the importance of variables

Figure 3 shows the results of the estimation of the LASSO regression coefficients for the explanatory variables in the prediction of the water quality index of the Rimac basin - Final WQI. Lags were added for each variable in order to capture the existing time dependence and to identify the variables and lags of greater importance in the prediction of the Final WQI. Also, to reduce the risk of overfitting and to facilitate the interpretability of the estimates. Finally, it is identified that the current and past values of biochemical oxygen demand (BOD), phosphate and pH are the parameters that have the greatest impact on the variability of Final WQI. Likewise, the results are consistent with the pollution dynamics of the Rimac River.

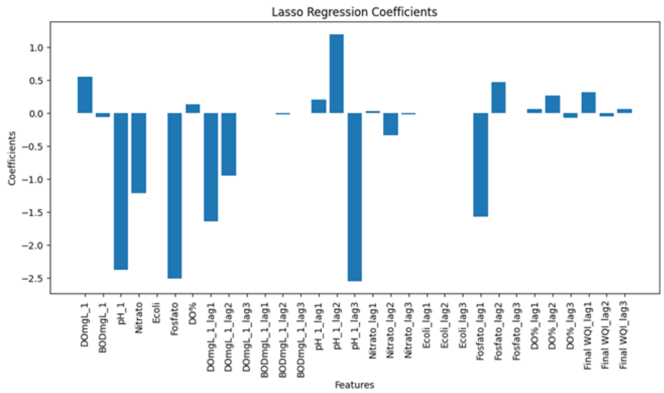

Fig. 3 Importancia de variables mediante los coeficientes asignados por el modelo de regresión LASSO

C. Descriptive analysis of the target variable Final WQI

TABLE III

FINAL WQI STATISTICS

Statistic	Value
Mean	64.491446
Std	9.345115
Min	37.650016
25%	59.384340
50%	65.819789
75%	70.409148
Max	81.543089

Table III shows a mean of 64.49 for the target variable - Final WQI, with a dispersion of 9.35, indicating moderate variability of the dataset with respect to its central tendency measure, the mean. The minimum value of the watershed water quality index is 37.65 and the maximum reaches the value of 81.54. 25% of the basin's water quality index values are below 59.38, the median is 65.82, while 75% of the water quality index values are below 70.41. Thus, the initial analysis concludes that the target variable Final WQI has a left-skewed distribution with the presence of extreme values.

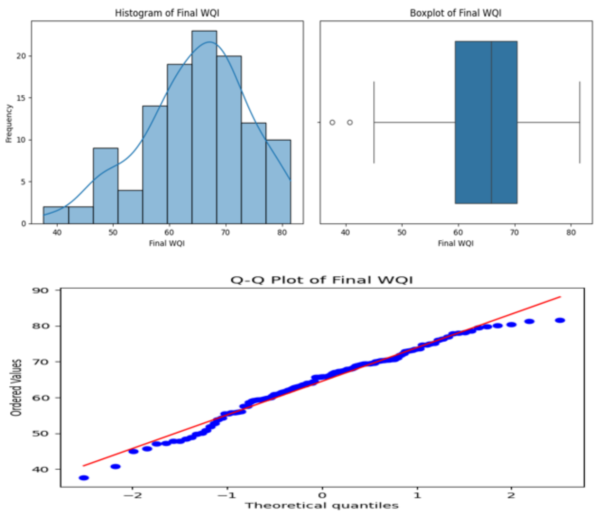

Fig. 4 Descriptive analysis of the final variable WQI

The descriptive analysis of the target variable in Figure 4 strengthens the evidence from Table III through the histogram, box plot, and quantile-quantile plot. The histogram shows a left-skewed unimodal distribution with a range from approximately 40 to 80, with a higher concentration between 60 and 70. Similarly, the box plot confirms the histogram analysis of a high concentration of the target variable Final WQI values between 60 and 70 and highlights the presence of outliers in the lower range, around 40-45. Therefore, the graphical analysis suggests a lack of normality in its distribution. Similarly, the quantile-quantile plot confirms the absence of normality in its distribution.

D. *Time series analysis: Final WQI*

Table IV presents evidence of stationarity of the target variable Final WQI based on the Augmented Dickey-Fuller (ADF) test. The results of this test show a p-value lower than the 5% significance level (0.000096 < 0.05), which allows us to reject the null hypothesis that the target variable is not stationary. This suggests that the mean and variance of the Final WQI series are constant over time, a neutral characteristic for the application of machine learning and deep learning techniques.

H0: The time series is not stationary.
H1: The time series is stationary.

TABLE IV
STATIONARITY TEST

Test	ADF statistic	P-value
the augmented dickey-fuller	-4.669995	0.000096

Table V shows the results of the Ljung-Box autocorrelation test for the target variable Final WQI (Water Quality Index). The results indicate a p-value significantly lower than the 5% significance level (1.051134e-09 < 0.05), which provides sufficient statistical evidence to reject the null hypothesis of absence of autocorrelation. Therefore, the test results indicate that the time series of the Final WQI variable exhibits significant autocorrelation, meaning that the observations are correlated over time. This autocorrelation is essential for the internal structure of the target variable Final WQI.

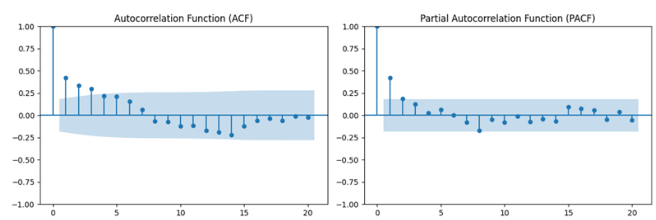

Fig. 5 Final WQI correlogram.

Fig.5, presents the correlograms of the autocorrelation functions (ACF) and partial autocorrelation functions (PACF); visually, the analysis strengthens the presence of autocorrelation in the Final WQI variable. In a study in Mexico, similar techniques were used to identify the nature of the target variable series [12].

TABLE V
AUTOCORRELATION TEST

test	lb_stat	lb_pvalue
ljung-box hypothesis test	62.83174	1.051134e-09

Thus, the results of the autocorrelation and stationarity tests in Figure 5 and the statistical tests in Tables IV and V provide a solid foundation for selecting the XGBoost algorithm as the meta-model in the stacking ensemble. This selection is suitable for capturing the stationary nature and temporal dependencies of the series, which will allow for more accurate and reliable predictions of the water quality index. Given that the analysis results show evidence that the target variable Final WQI is stationary and exhibits temporal dependence, it is concluded that there is a dependency structure in the series that needs to be modeled. These characteristics justify the identification of a predictive approach that is suitable for capturing both stationarity and temporal dependence. Therefore, it was decided to use the XGBoost algorithm as a meta-model in the stacking ensemble method for predicting the Final WQI variable.

The use of XGBoost as a meta-model in the stacking ensemble method allows for the combination of multiple base models, each specialized in capturing different aspects of the time series, such as autocorrelation, non-linearity, and complexity in its dynamics. This integrative approach enhances the model's generalization capacity and optimizes its predictive performance by combining the results of the base algorithms: decision tree, random forest, and multilayer perceptron neural network. Given that the time series is stationary and exhibits temporal dependence, the use of machine learning such as stacking with XGBoost offers an optimal solution for the precise prediction of the target variable Final WQI.

TABLE VI

HYPERPARAMETERS OF THE STACKING ASSEMBLY MODEL.

Role	Algorithm	Hyperparameters	Python package	Values
Modelo base	Decision tree	criterion, max_depth	scikit-learn	"gini", 10
Modelo base	Random forests	n_estimators, max_depth	scikit-learn	100, 20
Modelo base	Multi-Layer Perceptron	number of hidden layers, optimizer	scikit-learn	4, "adam"
Modelo meta	XGBoost	n_estimators, learning_rate	xgboost	100, 0.1

The structure of the ensemble stacking model used in the process of predicting the water quality of the Rimac River is shown in Table VI, the final configuration of the hyperparameters, the base models and the metamodel. The base algorithms selected are Decision Tree, Random Forest and Multilayer Perceptron, being Decision Tree and Random Forest of a bagging nature with the objective of reducing random error and Multilayer Perceptron to capture the complexity of the Final WQI target variable while the metamodel used is XGBoost due to the finding of stationarity in the Final WQI output with the objective of reducing systematic error. For each of these algorithms, two of their key hyperparameters and their final experiment settings are identified along with the Python packages used.

E. Results of the Stacking Ensemble predictive model with heterogeneous algorithms

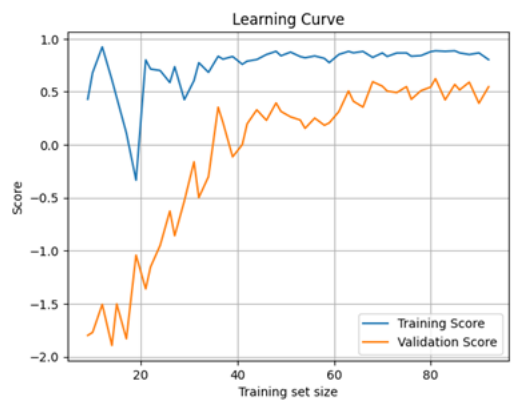

Fig. 6. Learning curve of the ensemble stacking model for the prediction of the water quality index Final WQI

Fig. 6 presents the learning curve of the model applied for the prediction of the target variable Final WQI, which is a continuous quantitative variable. Likewise, in the training (blue line), it is observed that the model has a low bias, as it quickly achieves high performance with a score close to 0.8. On the other hand, in the validation (orange line), an erratic behavior is initially observed, with negative scores, which indicates underfitting at the beginning due to the sample size in the training. As the size of the training set increases, the validation curve shows significant improvements, indicating that the model is better fitting the underlying patterns of the Final WQI variable. Thus, the difference between the training and validation curves is decreasing as the data increases, which suggests that the model does not exhibit a high degree of variation. However, the gap between both curves suggests that the model could still benefit from more data to reduce the variability between training and validation.

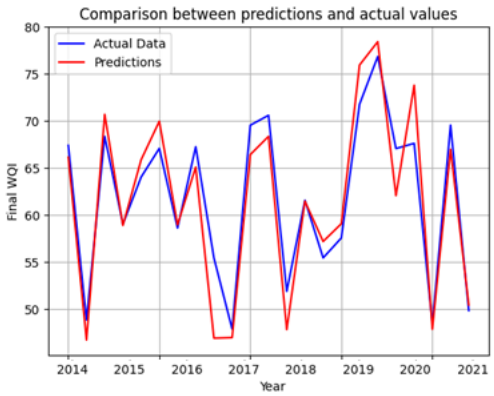

Fig. 7. Prediction model adjustment level
n of the water quality index Final WQI

Fig. 7 presents the performance of the water quality index prediction model (Final WQI). On the horizontal axis, the years (2014-2021) are represented, while on the vertical axis, the value of the Final WQI is shown. A good level of fit of the model used with respect to the empirical evidence is evident. The selected model combines the predictions of multiple base algorithms with a metamodel to reduce random error. In this case, the results suggest that the model satisfactorily captures the fluctuations and general trends of the target variable during the analyzed period, with minor deviations at some specific points. This type of graphical evaluation is common in machine learning, as it allows for validating the model's generalization capability and detecting possible limitations in its performance.

TABLE VII
PREDICTION ERROR METRICS AND GOODNESS OF FIT.

Model	MSE	MAE	R^2
Ensamble Stacking Model	9.954	2.433	0.859

The results in Table VII demonstrate a good performance of the predictive model for the water quality index (Final WQI), with a mean square error of 9.95 and a mean absolute error of 2.43, indicating low errors in the predictions. Likewise, a coefficient of determination of 0.859 is shown, meaning that the stacking ensemble method captures 85.9% of the total variability of the data, effectively explaining the general trends.

IV. CONCLUSIONS

The results of the analysis of the physicochemical and microbiological data of the Rímac basin and the Chillón-Rímac-Lurín jurisdiction (2014-2021) allowed to propose a predictive model of the water quality index (NSF WQI) using the ensemble stacking method, identifying the XGBoost algorithm as a metamodel based on the results of the Augmented Dickey-Fuller test ($0.000096 < 0.05$) that show stationarity. Likewise, evidence of temporal dependence is presented through the analysis of the correlograms and the Ljung-Box autocorrelation tests ($1.051134e-09 < 0.05$) in the target variable (NSF WQI).

The prediction model showed robust performance with a coefficient of determination (R^2) of 85.9%. Furthermore, the Mean Square Error (MSE) of 9.9 and the Mean Absolute Error (MAE) of 2.4 show a low level of error in the predictions, with reduced average deviations between the predicted and actual values.

The proposed model showed slight signs of overfitting, indicating that an increase in sample size improves the model's generalization. However, the results obtained confirm that the stacking ensemble model with heterogeneous algorithms is an effective tool for forecasting the water quality index – Final WQI, allowing for precise estimates.

ACKNOWLEDGMENT

The author expresses his gratitude to God and to the research direction of the Lima Region of the Universidad Tecnologica del Peru (R+D+i Project Competition, UTP 2023-1).

REFERENCES

[1] Behar, R., Zúniga de Cardozo, Ma. C. y Rojas, O. “Análisis y valoración del índice de calidad de agua (ICA) de la NSF. casos ríos Cali y Meléndez”. Revista Ingeniería y Competitividad, 1 (1): 17-27, 1997

[2] [2] Zubaidah T., Hamzani S., Arifin A.”analyzing the impact of dissolved organic components on river water quality and its implications for human health: a case study from banjar district”, Jurnal Kesehatan Lingkungan, 16 (2), pp. 181 - 189, (2024), doi: 10.20473/jkl.v16i2.2024.181-189

[3] Sharma M., Rawat S., Kumar D., Awasthi A., Sarkar A., Sidola A., Choudhury T., Kotecha K.” The state of the Yamuna River: a detailed review of water quality assessment across the entire course in India”, Applied Water Science, 14 (8), art. no. 175, (2024), doi: 10.1007/s13201-024-02227-x

[4] Latif M.B., Khalifa M.A.K., Hoque M.M.M., Ahammed M.S., Islam A., Kabir M.H., Tusher T.R. ” Appraisal of surface water quality in vicinity of industrial areas and associated ecological and human health risks: a study on the Bangshi river in Bangladesh”, Toxin Reviews, 41 (4), pp. 1148 - 1162, (2022), doi: 10.1080/15569543.2021.1978498

[5] Nair J.P., Vijaya M.S. ” River Water Quality Prediction and index classification using Machine Learning”, Journal of Physics: Conference Series, 2325 (1), art. no. 012011, (2022), doi: 10.1088/1742-6596/2325/1/012011

[6] Zhang H., Ren X., Chen S., Xie G., Hu Y., Gao D., Tian X., Xiao J., Wang H. “Deep optimization of water quality index and positive matrix factorization models for water quality evaluation and pollution source apportionment using a random forest model”, Environmental Pollution, 347, art. no. 123771, (2024), doi: 10.1016/j.envpol.2024.123771

[7] Masteali S.H., Bayat M., Bettinger P., Ghorbanpour M., “Uncertainty analysis of linear and non-linear regression models in the modeling of water quality in the Caspian Sea basin: Application of Monte-Carlo method”, Ecological Indicators, 170, art. no. 112979, 2025

[8] Uypatchawong S., Chanamarn N., “Enhancing surface water quality prediction efficiency in northeastern thailand using machine learning,” *Indonesian Journal of Electrical Engineering and Computer Science, 36 (2), pp. 1189 - 1198,* 2024.

[9] Khang A., Hajimahmud V.A., Singh K., “Water quality classification using machine learning algorithms”, Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, pp. 60 - 76, 2024

[10] Najafzadeh M., Ghaemi A., “Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods,” Environmental Monitoring and Assessment, 191 (6), art. no. 380, 2019

[11] Haribowo R., Dermawan V., Fitrina H., “A study of water quality modelling with the artificial neural network method in Surabaya river,” IOP Conference Series: Earth and Environmental Science, 437 (1), art. no. 012003, 2020

[12] del Castillo A.F., Garibay M.V., Díaz-Vázquez D., Yebra-Montes C., Brown L.E., Johnson A., Garcia-Gonzalez A., Gradilla-Hernández M.S, “Improving river water quality prediction with hybrid machine learning and temporal analysis,” , *Ecological Informatics, 82, art. no. 102655, 2024*.

[13] Noori N., Kalin L., Isik S., “Water quality prediction using SWAT-ANN coupled approach,” , *Journal of Hydrology, 590, art. no. 125220, 2020*

[14] Lu Y.-M., Zhang H., Shao W., “Prediction of river water quality considering spatiotemporal correlation and meteorological factors,” , *9th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2021*

[15] Satish N., Anmala J., Rajitha K., Varma M.R.R.,“ Prediction of stream water quality in Godavari River Basin, India using statistical and artificial neural network models” H2Open Journal, 5 (4), pp. 621 - 641, *2022*

[16] Liu H., Li D., Liu C., Fang B., He W., “Multi-station Water Quality Prediction Considering Temporal and Spatial Correlation: Based On T-GCN” , *ACM International Conference Proceeding Series, art. no. 22, 2023*

[17] Uğurluoğlu A., Uyguçgil H., “Assessing performance of spatial statistical stream network models in predicting alkalinity: A comparative study of two river basins in Türkiye”, *Ecological Informatics, 78, art. no. 102358, 2023*

[18] Fonseca A., Botelho C., Boaventura R.A.R., Vilar V.J.P., “Evaluating the Effects of Parameter Uncertainty on River Water Quality Predictions” , *Resources, 13 (8), art. no. 106, 2024*

[19] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, 1992.

[20] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Belmont, CA: Wadsworth, 1984.

[21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[22] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[23] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[24] Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)? —Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7, 1247–1250, 2014.