International Journal of Data Science and Analysis
2025; 11(6): 171-177
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20251106.12

SciencePG
Science Publishing Group

Stacked Ensemble Classifier for Adoption of Point-of-Collection Water Treatment Technology Among Households in Western Kenya

Robert Omondi Oluoch[1, *], Herbert Imboga[1], Anthony Waititu[1], Susan Mwelu[1], Illah Evance[2], Ferdnand Ongeta[2]

[1]Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya
[2]Evidence Action, Africa Regional Office, Nairobi, Kenya

Email address:

robertdon388@gmail.com (Robert Omondi Oluoch), awaititu@jkuat.ac.ke (Anthony Waititu), imbogaherbert@jkuat.ac.ke (Herbert Imboga), smwelu@jkuat.ac.ke (Susan Mwelu), Illah.evance@evidenceaction.org (Illah Evance), ferdnard.ongeta@evidenceaction.org (Ferdnand Ongeta)
*Corresponding author

To cite this article:

Robert Omondi Oluoch, Herbert Imboga, Anthony Waititu, Susan Mwelu, Illah Evance, Ferdnand Ongeta. (2025). Stacked Ensemble Classifier for Adoption of Point-of-Collection Water Treatment Technology Among Households in Western Kenya. *International Journal of Data Science and Analysis, 11*(6), 171-177. https://doi.org/10.11648/j.ijdsa.20251106.12

Received: 17 September 2025 ; **Accepted:** 29 September 2025 ; **Published:** 10 November 2025

Abstract: The Dispensers for Safe Water program under Evidence Action promotes point of collection water treatment through the installation of chlorine dispenser gadgets in rural parts of Kenya. Although the initiative has improved access to safe drinking water, monitoring household adoption remained a challenge during the COVID-19 pandemic, which limited field-based data collection and led to increased dependence on phone surveys. In addition, technology adoption data are often imbalanced, which poses difficulties for traditional classification methods. This study aimed to develop and implement a stacking ensemble classifier to model the adoption of chlorine dispensers among households in western Kenya. Data were collected from 27,457 households. The analysis used structured household, promoter and spot check survey data. The key variables included chlorine availability, user knowledge, household demographics, and engagement with promoters. RF, ANN, and NB models were trained and evaluated individually, then combined using a stacked ensemble approach. The ensemble model outperformed all base learners, achieving the highest accuracy (69.1%) and AUC (0.6959). The variable importance analysis revealed that the presence of chlorine and the knowledge of the user were the strongest predictors of adoption. In conclusion, ensemble learning provides a reliable method for modeling behavioral adoption in public health interventions. The findings offer practical insights for programs and demonstrate the potential of machine learning in improving, targeting and monitoring of safe water initiatives in low-resource settings.

Keywords: Stacked Ensemble, Random Forest, Artificial Neural Networks, Naive Bayes, Total Chlorine Residue, Machine Learning

1. Introduction

Access to safe drinking water is essential for public health, particularly in rural communities of developing countries, where waterborne diseases such as diarrhea remain a prominent cause of morbidity and mortality among children. According to the World Health Organization, unsafe water, sanitation, and hygiene contribute to millions of deaths annually, with the burden disproportionately affecting developing regions [12].

In response, Evidence Action's Dispensers for Safe Water (DSW) initiative installs chlorine dispensers at communal water points, which provides convenient and cost effective access to safe drinking water. By 2022, the program had reached over 4.5 million people in Kenya, Uganda, and Malawi [1]. The program's success is hugely dependent on community adoption, supported by local promoters who maintain the dispensers and encourage households on the usage.

The COVID-19 pandemic disrupted in-person monitoring activities, which limited the direct measurement of chlorine presence in the drinking water. The program shifted to phone-based surveys, introducing challenges of self-report bias on adoption and imbalanced datasets, where adopters and non-adopters are rarely evenly distributed. The imbalance reduce the accuracy of traditional classification models, which in most cases favor the majority class.

This study addressed these challenges by developing a stacked ensemble classifier that combined Random Forest, Artificial Neural Network, and Naive Bayes. The model aimed to better the prediction of adoption patterns, hence strengthen monitoring, and guide more targeted interventions for the DSW program, which ultimately enhance its impact on safe water access.

2. Literature Review

Access to safe water substantially reduces diarrheal diseases which is a significant public health challenge in Kenya [9, 15]. Waterborne pathogens, such as diarrheagenic Escherichia coli, contribute to high morbidity, especially among under five [5]. The chlorine dispensers technology offer a cost-effective solution, reducing microbial contamination [10]. Adoption of the technology varies based on the technology awareness, infrastructure functionality, and household practices [4].

Imbalanced datasets which is a common occurance in technology adoption studies, pose challenges for traditional classifiers, which favor the majority class [7]. Machine learning offers robust solutions for public health, with applications in disease prediction and behavioral modeling [2, 8, 11]. Models driven by data and AI algorithms enhance the identification of relevant features and contribute to more robust modelling and prediction frameworks [6].

Ensemble techniques such as stacking, combine multiple models to improve its predictive performance [3]. Study conducted by [16] demonstrated the efficacy of stacked ensembles in health related classification tasks, making them suitable for modeling new technology adoption.

3. Materials and Methods

3.1. Study Site and Population

The study involved data from nine counties in Western Kenya (Bungoma, Busia, Homa Bay, Kakamega, Migori, Siaya, Trans Nzoia, Uasin Gishu, and Vihiga), where the DSW program operates. The counties were selected because they represent diverse rural communities where chlorine dispensers have been installed at communal water points. The study population consisted of households that regularly use the water sources installed chlorine dispensers gadgets.

The study involved 27,457 households. Data collection involved providing responses through structured questionnaires and spot-check observations. It captured household demographics, water collection and storage practices, promoter engagement, dispenser status, and adoption behavior.

Figure 1. *Map of Kenya: The study area highlighted in Blue*

3.2. Study Design and Variables

Quantitative research design was utilized to model the household adoption of point of collection water treatment technology. Adoption was defined as the presence of detectable chlorine residuals in household drinking water samples.

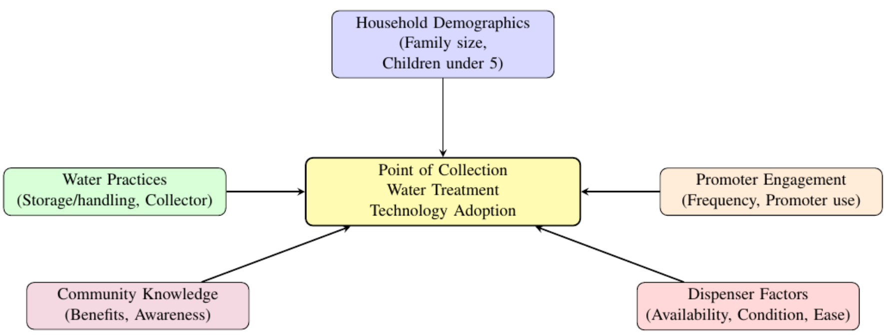

Figure 2. *Study Variables.*

3.3. Sample Size and Sampling

The required sample size was determined using Cochran's formula as presented below.

$$n = \frac{Z^2 \cdot p \cdot (1-p)}{e^2}$$

where $Z = 1.96$ (95% confidence), $p = 0.5$ (maximum variability), and $e = 0.05$ (margin of error).

A multi-stage clustered sampling method was employed, where water points with installed dispensers were randomly selected within each county. From each water point, one promoter and four households were selected, ensuring representativeness across the nine counties.

3.4. Data Collection Tools

The study used three structured data collection instruments as outlined below.

1. *Household Questionnaire* - collected data on socio-demographics, water treatment practices, and promoter engagement.
2. *Promoter Questionnaire* - collected data on the promoter characteristics, chlorine use, and reported challenges.
3. *Spot-Check Observation Form* - collected data on the physical condition and functionality of chlorine dispensers and water sources.

This triangulated approach ensured standardized, reliable, and complementary data.

3.5. Data Analysis and Modeling

Data analysis was conducted using R software, with focus on the development of a stacked ensemble classifier combining RF, ANN, and Naive Bayes.

The modeling process involved the following stages.

1. *Data Normalization and Preparation:* Data point eatures were standardized as

$$x'_j = \frac{x_j - \mu_j}{\sigma_j}$$

to ensure numerical stability. Data were split into training (80%) and test (20%) sets.
2. *Model Training:* The base models (RF, ANN, NB) were trained using the training dataset, with hyperparameters tuned through cross-validation where k = 5.
3. *Stacked Ensemble:* Predictions from base models were concatenated into a new dataset and used as inputs for a Random Forest meta-model.
4. *Model Evaluation:* Performance was assessed using accuracy, precision, recall, and AUC-ROC. Cross-validation technique ensured generalizability and robustness of the modelling results.

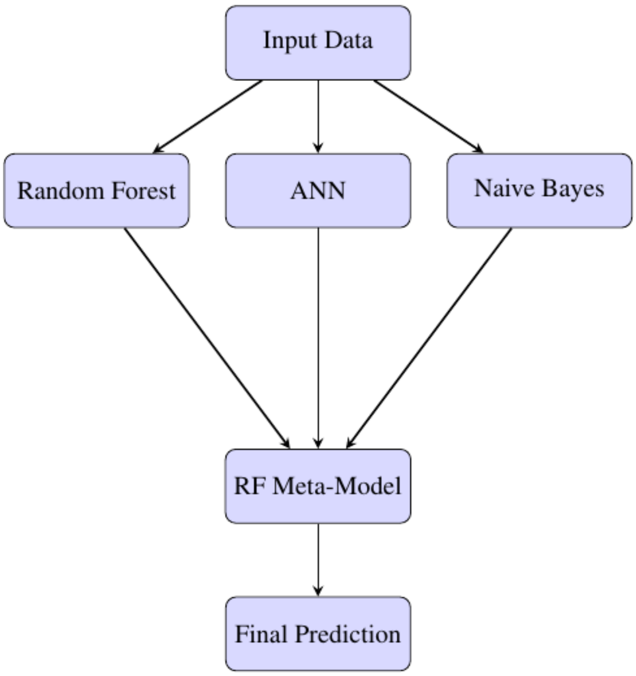

Figure 3. *Stacked Ensemble Model Framework.*

3.6. Model Specifications

To model households adoption of chlorine dispensers technology, ANN, RF, NB and Stacked Ensemble models were employed.

3.6.1. Artificial Neural Networks (ANN)

The study utilized the ANN model that was structured with an input layer comprising of p predictor variables $\{x_1, x_2, \ldots, x_p\}$, followed by hidden layers composed of h_j which applied the nonlinear activation (ReLU), and a softmax output layer. Training was conducted by minimizing the crossentropy loss, using backpropagation technique with gradient descent. [17].

3.6.2. Naive Bayes (NB)

The Naive Bayes classifier was implemented on the basis of Bayes' theorem, under the simplifying assumption that predictor variables $\{x_1, x_2, \ldots, x_p\}$ are conditionally independent given the outcome class C_k. Considering continuous predictors x_j, the model assumed a Gaussian distribution, while categorical predictors were modeled using frequency-based probabilities. Estimation of Class prior probabilities $P(C_k)$ were directly from the training data, and prediction was made by selecting the class with the highest posterior probability $\hat{y} = \arg\max_k P(C_k \mid \mathbf{x})$. [18].

3.6.3. Random Forest (RF)

The Random Forest algorithm was applied in this study as an ensemble of B decision trees, where each tree was trained on a bootstrap sample D_b originating from the training data. At each split, a random subset of m predictors was considered from the original set of $\{x_1, x_2, \ldots, x_p\}$, which in return helped to reduce correlation among trees and improve generalization in the model. The final model prediction $\hat{y}_i$ for each observation i was achieved through majority voting across all the trees in the ensemble $\hat{y}_i = \arg\max_y \sum_{b=1}^{B} I(\hat{y}_i^b = y)$, where $\hat{y}_i^b$ denotes the prediction of the b-th tree [14].

3.6.4. Stacked Ensemble

Stacked ensemble model was constructed to combine the predictive strengths of ANN, NB, and RF models. The class probability outputs from the base learners formed the meta-feature matrix formulated as follows

$$\mathbf{Z} = [\mathbf{P}_{ANN}, \mathbf{P}_{NB}, \mathbf{P}_{RF}] \tag{1}$$

A Random Forest was employed as the meta-learner, trained on $\mathbf{Z}$. The final ensemble prediction was derived as follows

$$\hat{y} = \arg\max_{y} P(y \mid \mathbf{Z}) \tag{2}$$

4. Results and Discussion

4.1. Descriptive Statistics

The analysis results showed that 44% of households adopted chlorine treatment, 74% reported to know dispenser gadget usage steps, similarly, 74% reported to know waiting time after treatment. Infrastructure was robust, with 99% usable water points and 100% dispenser presence at the water point. Chlorine availability (86%) and dispensing (87%) were high, with only 8% reporting dispenser hardware issues.

Table 1. *Summary Statistics for Key Variables.*

Variable	N	Freq	Mean
TCR Adoption (+CL)	27,457	15,268	44%
TCR Adoption (-CL)	27,457	12,189	56%
Promoter Interaction (Past 30 Days)	27,457	8,307	30%
Knows to Wait 30 Min	27,457	20,184	74%
Knows Treatment Steps	27,457	20,262	74%
Knowledge Score	27,457	5,997	22%
Good Storage Practice	26,494	7,519	28%
Water Point Usable	27,457	27,184	99%
Dispenser Present	27,457	27,457	100%
Chlorine in Tank	27,457	23,489	86%
Chlorine Dispensed	27,457	23,984	87%
Hardware Problem	27,457	2,161	8%
Household Size	27,453	–	5
Child Under 5	27,453	–	1

4.2. Model Performance

The stacked ensemble model outperformed the base models on the independent test dataset. The stacked model achieved an accuracy of 69.1% and AUC of 0.6959. RF model performed well in sensitivity (74.1%), NB model excelled in identifying adopters (82.9% sensitivity), and ANN model showed balanced performance. McNemar's test confirmed that the ensemble's superior performance was not by chance $p < 0.05$ [13].

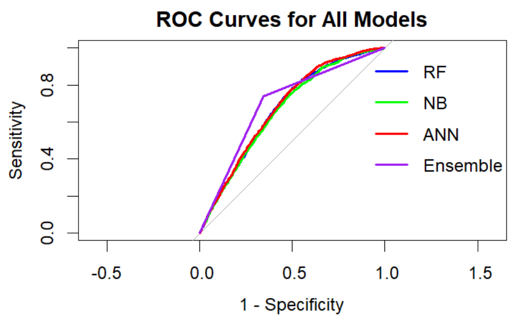

Figure 4. *ROC Curves for Each Model.*

Table 2. *Model Performance Metrics.*

Model	Accuracy	Kappa	Sensitivity	Specificity	Balanced Accuracy	AUC
Random Forest	0.6260	0.2663	0.7405	0.5345	0.6375	0.6707
Naive Bayes	0.5910	0.2174	0.8296	0.4005	0.6151	0.6633
Artificial Neural Network	0.6257	0.2570	0.6791	0.5831	0.6311	0.6769
Stacked Ensemble	0.6913	0.3852	0.6551	0.7367	0.6959	0.6959

4.3. Variable Importance

The RF model highlighted *Chlorine in Tank*, *Knowledge Score*, and *Chlorine Dispensed* as top predictors of adoption, underscoring the importance of chlorine supply availability and dispenser user knowledge.

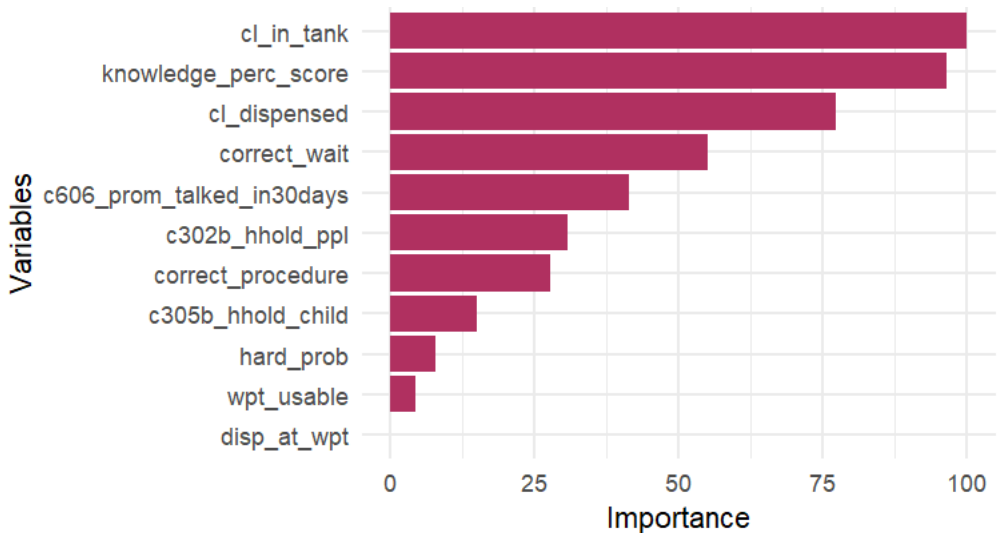

Figure 5. *Variable Importance Plot (Random Forest).*

5. Conclusion and Recommendations

5.1. Conclusion

The stacked ensemble classifier effectively modeled adoption, achieving superior performance (69.1% accuracy, 0.6959 AUC) compared to the individual base models, RF (62.6%, 0.6707), ANN (62.6%, 0.6769), and NB (59.1%, 0.6633). Key predictors of point of collection water treatment technology included chlorine availability and knowledge, aligning with public health literature [4]. The model's robustness supports its use in monitoring and targeting interventions for the DSW program.

5.2. Recommendations

Based on the results of this study, it is recommended that the stacked ensemble classifier be considered for practical use in modeling the point of collection water treatment technology. Integrating the model into program monitoring tools could help improve targeting and decision making. The ensemble model can be embedded into monitoring tools like dashboards, survey systems or mobile apps to automatically generate adoption risk scores immediately the new data are submitted as represented in *Figure 5*. These predictions help managers target resources to low-adoption communities and guide decisions on promoter visits or sensitization campaigns. Over time, integration supports adaptive management by tracking whether interventions close adoption gaps.

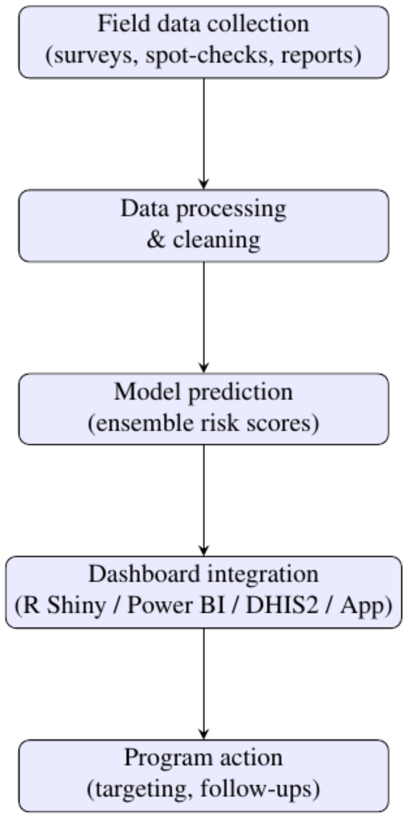

Figure 6. *Workflow for integrating the stacked ensemble into program monitoring tools (downward flow).*

To keep the model accurate and relevant, it should be regularly updated with new data. Future studies may also benefit from including more behavioral and contextual variables to improve prediction. Given the model's strong performance, similar ensemble approaches could be explored in other development areas where understanding household behavior is important.

ORCID

0009-0003-8714-0080 (Robert Omondi Oluoch)
0009-0003-9963-4977 (Herbert Imboga)
0000-0003-0268-2968 (Anthony Waititu)
0009-0005-9570-9112 (Susan Mwelu)
0009-0008-9949-5424 (Illah Evance)
0009-0001-6782-3365 (Ferdnand Ongeta)

Abbreviations

RF	Random Forest
ANN	Artificial Neural Network
NB	Naive Bayes
AUC	Area Under the Curve
DSW	Dispensers for Safe Water
TCR	Total Chlorine Residue
ROC	Receiver Operating Characteristic
JKUAT	Jomo Kenyatta University of Agriculture
WHO	World Health Organization

Acknowledgments

I am deeply grateful to my supervisors for their invaluable guidance, constructive feedback, and encouragement, which greatly enriched this study. I sincerely thank Evidence Action for providing the data and context through the Dispensers for Safe Water program. I also appreciate the support and camaraderie of my colleagues and fellow students at Jomo Kenyatta University of Agriculture and Technology, whose discussions broadened my perspective. Above all, I extend my heartfelt thanks to my family and friends for their unwavering support, understanding, and encouragement throughout this journey.

Author Contributions

Robert Omondi: Conceptualization, Formal Analysis, Investigation, Software, Validation, Visualization, Writing - original draft, Writing - review & editing

Herbert Imboga: Supervision, Writing - review & editing

Anthony Waititu: Supervision, Writing - review & editing

Susan Mwelu: Supervision, Writing - review & editing

Illah Evance: Methodology, Supervision, Writing - review & editing

Ferdnand Ongeta: Methodology, Supervision, Writing - review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Millennium Water Alliance. Learning from Piloting Dispensers for Safe Water. In: Washington, DC (2019).

[2] Meriem Amrane et al. Breast cancer classification using machine learning. In: 2018 Electric Electronics, Computer Science, Biomedical Engineering Meeting (EBBT) (2018), pp. 1-4. https://doi.org/10.1109/EBBT.2018.8391453

[3] Thomas G. Dietterich. Ensemble Methods in Machine Learning. In: International Workshop on Multiple Classifier Systems. Berlin, Heidelberg: Springer Berlin Heidelberg, June 2000, pp. 1-15. https://doi.org/10.1007/3-540-45014-9

[4] M. Husein et al. Examination of water quality at the household level and its association with diarrhea among young children in Ghana: Results from UNICEF-MICS6 survey. In: PLOS Water (2023). https://doi.org/10.1371/journal.pwat.0000049

[5] Y. Iijima et al. High Prevalence of Diarrheagenic Escherichia coli among Children with Diarrhea in Kenya. In: Japanese Journal of Infectious Diseases 70.1 (2017), pp. 80-83. https://doi.org/10.7883/yoken.JJID.2016.064

[6] H. Jallow et al. Transfer learning for predicting gross domestic product growth based on remittance inflows using RNN-LSTM hybrid model: a case study of The Gambia. In: Frontiers in Artificial Intelligence 8 (2025), p. 1510341. https://doi.org/10.3389/frai.2025.1510341

[7] Jinseok Kim and Jenna Kim. The impact of imbalanced training data on machine learning for author name disambiguation. In: Scientometrics 117.1 (2018), pp. 511-526. https://doi.org/10.1007/s11192-018-2865-9

[8] Sammy Kiprop et al. Classification of Contraceptive Use Among Undergraduate Students Using a Supervised Machine Learning Technique. In: American Journal of Theoretical and Applied Statistics 12.5 (2023), pp. 120-128. https://doi.org/10.11648/j.ajtas.20231205.14

[9] Dian Kurniawati et al. Poor Basic Sanitation Impact on Diarrhea Cases in Toddlers. In: Journal Name 13 (2021), pp. 41-47. https://doi.org/10.20473/JKL.V13I1.2021.41-47

[10] Daniele Lantagne. Sodium hypochlorite dosage for household and emergency water treatment. In: Journal American Water Works Association 100 (Aug. 2008), pp. 106119. https://doi.org/10.1002/j.1551-8833.2008.tb09704.x

[11] S. Mooney and V. Pejaver. Big Data in Public Health: Terminology, Machine Learning, and Privacy.

In: Annual Review of Public Health 39 (2018), pp. 95-112. https://doi.org/10.1146/annurev-publhealth-040617-014208

[12] K. J. Nath, Sally Bloomfield, and Martin Jones. Household water storage, handling and point-of-use treatment. In: 2006. https://api.semanticscholar.org/CorpusID:75277027

[13] Matilda Q. R. Pembury Smith and Graeme D. Ruxton. Effective use of the McNemar test. In: Behavioral Ecology and Sociobiology 74.11 (2020), p. 133. https://doi.org/10.1007/s00265-020-02916-y

[14] Hasan Ahmed Salman, Ali Kalakech, and Amani Steiti. Random Forest Algorithm Overview. In: Babylonian Journal of Machine Learning 2024 (2024), pp. 69-79. https://doi.org/10.58496/BJML/2024/007

[15] K. Ásia Chaves da Silva and C. N. K. Suda. Acute diarrheal diseases and their relationship with water quality in Araguatins, Tocantins, Brazil: a cross-sectional study. In: Bioscience Journal (2023). https://doi.org/10.14393/bj v39n0a2023-62592

[16] Bayu Adhi Tama, Sun Im, and Seungchul Lee. Improving an Intelligent Detection System for Coronary Heart Disease Using a Two-Tier Classifier Ensemble. In: BioMed Research International (2020). https://doi.org/10.1155/2020/9816142

[17] Yu-chen Wu and Jun-wen Feng. Development and Application of Artificial Neural Network. In: Wireless Personal Communications 102.2 (2018), pp. 1645-1656. https://doi.org/10.1007/s11277-017-5224-x

[18] Feng-Jen Yang. An Implementation of Naive Bayes Classifier. In: 2018 International Conference on Computational Science and Computational Intelligence (CSCI). 2018, pp. 301-306. https://doi.org/10.1109/CSCI46756.2018.00065