STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **11**, January 2023, pp 70–81.
Published online in International Academic Press (www.IAPress.org)

Indicating if water is safe for human consumption using an enhanced machine learning approach

Nachaoui Mourad[1], Lyaqini Soufiane[2], Chaouch Marouane[1]*

[1]*Equipe de Mathématiques et Intercations, Université Sultan Moulay Slimane, Béni-Mellal, Morocco*
[2]*Hassan First University of Settat, Ecole Nationale des Sciences Appliquees, LAMSAD Laboratory, Morocco*

Abstract Predicting water quality accurately is critically important in real-life water resource management. This work proposes an approach based on supervised machine learning to predict water quality. Motivated, by the success of the non-smooth loss function for supervised learning problems [22], we reformulate the learning problem as a regularized optimization problem whose fidelity term is the hinge loss function and the hypothesis space is a polynomial approximation. To deal with the non-differentiability of the loss function, a special smoothing function is proposed. Then, the obtained optimization problem is solved by an improved conjugate gradient algorithm. Finally, some experiments results are presented.

Keywords Supervised learning, Smooth approximation, Hing loss, Tikhonov regularization, Taylor polynomials, Conjugate gradient, Water quality

AMS 2010 subject classifications 68T05; 90C90

DOI: 10.19139/soic-2310-5070-1703

1. Introduction

Water is the life source for human beings and all other living things. Thanks to this essential element, these living beings can survive, grow and develop. Indeed, water is unarguably an essential commodity after air with an unparalleled value and there is virtually nowhere water is not important. However, due to overpopulation, urbanization, and lack of proper water resource management, the available freshwater does not meet the demand resulting in water scarcity. These explain the clean water crisis that engulfs various countries in the world with only 1% of clean water that can be consumed by humans. With a very small amount of good clean water, it is difficult for residents to access clean water. As many as 663 million people who are sourced from WHO data show that it is difficult for them to access clean water [1]. Based on UNESCO data, by 2025 it is predicted that two-thirds of the world's population will live in areas that lack clean water [2]. The World Water Assessment Program (WWAP) has predicted the condition of clean water for the next few years, WWAP is under UNESCO. Under certain conditions, as an illustration of the need for water in daily life, not less than 85% of clean water turns into waste water. Everyone can use up to 100 liters of water per day to meet their daily needs [3]. Water and sanitation are closely related to clean water management because proper sanitation facilities can meet the needs of a clean and healthy life which is an important element in improving the health status of the world's population. In its application, drinking water is recommended to be drunk directly, it can be seen its content based on microbiological, physical chemistry and radio active substances [4, 5, 6]. Water in the human body ranges from 50-70% of all human body weight. The

*Correspondence to: Marouane Chaouch (Email: marouanechaouche481@gmail.com). Equipe de Mathématiques et Intercations, Université Sultan Moulay Slimane, Béni-Mellal, Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

human body contains water which consists of 80% water content in the blood, on the contrary, if a deficiency of 15% of body weight can result in death [7, 8, 9].

The description of water quality is usually described in the form of parameters and variables. Several various parameters are used as the basis for determining the model as in paper [10, 11, 12]. Based on these parameters, in several studies, it is used to determine the prediction of the next parameter value, which is stated in the paper [13].

The main contribution of this paper is the elaboration of a supervised learning approach for solving the prediction of water quality problems. In this case, the importance of the water quality features is studied. This is done through the determination of which covariates are important factors in water potability. Furthermore, based on our predictive model, the feature that's the foremost important in water quality is find out.

For the technique of resolution, the model is reformulated into a regularized optimization problem with the term of the fidelity is the nonsmooth loss function called Hinge function and the hypothesis space is constructed by a polynomial approximation. Indeed, polynomial approximation of functions is particularly important in the area of statistical learning theory because of the famous represented theorem which states that any function in a Hilbert space that minimizes an empirical risk function can be written as a polynomial approximation of the function evaluated at the training set. The use of nonsmooth loss function for supervised learning problem gives more consistent results [20, 21, 22, 23, 24, 25]. To address the difficulty of nondifferentiability, we introduce the smoothing technique [21, 22], which transforms the optimization problem into a smooth one. Furthermore, the fidelity term of the resulting optimization problem is twice differentiable and convex, which is solved directly using Tikhonov regularization and Conjugate Gradient algorithm [26]. Finally, the proposed approach is supported by several numerical validations and the efficiency is examined by comparison with the Support Vector Machine (SVM)[16], the Naive Bayes (NB) [19] and the K-Nearest Neighbours (KNN)[18] algorithms in terms of the accuracy and precision. Obtained experimental result on the water potability data sets indicates that the proposed approach is an accurate and helpful tool in machine learning.

The organization of this paper is as follows. In section 2, we present the setting problem and its reformulation as a minimization one, using Hinge loss function. Section 3 is concerned to a polynomial approximation of the model. Furthermore, a numerical algorithm based on Conjugate Gradient algorithm is presented. In section 4, we evaluate the efficiency of the proposed approach using water potability data sets and compare it with Support Vector Machine (SVM), Naive Bayes (NB), and K-Nearest Neighbours (KNN) algorithms.

2. Statement and formulating the supervised learning problem

The problem of supervised learning aims is to find the solution of the following minimization problem:

$$\min_{h_{\mathcal{T}} \in H} \mathcal{J}(h) := \frac{1}{n} \sum_{i=1}^{n} \ell(y_i, h(x_i)) + \lambda \|h\|^2 \tag{1}$$

where $h_{\mathcal{T}}$ is a function belonging to $\mathcal{H}$, the hypothesis space that is being searched(which we discuss in the next subsection), $\mathcal{T}$ a training set of size n, $\mathcal{T} = \{(x_i, y_i)\}_{i=1}^{n}$, and ℓ is a non-negative loss function measuring the degree of agreement between $h(x)$ and y. λ is the regularization parameter (in practice, it is defined using validation data), chosen to make sure of the discrepancy.

Our aim in this paper is to perform a binary classification i.e. to classify observations that take feature values into one of two classes, -1 or $+1$. A discriminant model h yields a classifier $sgn(h(x))$ that represents our guess of a target value for a future observation.

Throughout this paper we use the hinge loss given by:

$$\ell(\omega, \nu) = max(1 - \omega\nu, 0). \tag{2}$$

3. The Proposed Approach

3.1. Polynomial approximation

Polynomials are very simple mathematical functions that have the ability to represent very generally non-linear relations. The approximation of more complicated functions by polynomials is a basic element for many digital techniques.

The goal is to approach a function that is difficult to evaluate, such as a density or a distribution function, for the purpose of a quick evaluation of a computer. Here, there is no interest in the polynomial curve itself. Rather, the interest is to know how much the polynomial can follow the special function, and in particular the weakness of the maximum error. Very high order polynomials can be used here if they provide accurate approximations.

Since we're interested in defining a hypothesis space that is being searched, we've decided on approximating the hypotheses as polynomials. For this reason, we'll briefly introduce Taylor Polynomials.

Definition 1. Taylor Polynomial: For a given function $f(x)$ and $a \in \mathbb{R}$, the Taylor polynomial of degree n about $x = a$ is defined as

$$\begin{aligned} T_n(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \end{aligned} \tag{3}$$

Here we are assuming that f is differentiable n times at a.

In our case we'll be studying a multivariable problem, introducing Taylor's theorem for higher dimensions is necessary:

Definition 2. Taylor Polynomial– Multivariable Suppose that $S \subseteq \mathbb{R}^n$ is an open set and that $f : S \leftarrow \mathbb{R}$ is a function of class C^k on S. For a point $a \in S$, the k^{th} order Taylor polynomial of f at a is the unique polynomial of order at most k, denoted $T_{a,k}(x-a)$, such that

$$\begin{aligned} f(a) &= T_{a,k}(0), \\ \partial^\alpha f(a) &= \partial^\alpha T_{a,k}(0) \quad \text{for all partial derivatives of order up to k.} \end{aligned} \tag{4}$$

For completeness, we state the formula for the kth order Taylor polynomial, for arbitrary $k \in \mathbb{N}$. First, we introduce the notation:

$$\partial^\alpha f = (\frac{\partial}{\partial x_1})^{\alpha_1}(\frac{\partial}{\partial x_2})^{\alpha_2} \ldots (\frac{\partial}{\partial x_n})^{\alpha_n} \tag{5}$$

where α is a multi-index; that is α has the form $(\alpha_1, \ldots, \alpha_n)$, where each α_j is a non-negative integer. For such a multi-index, we will also use the notation

$$\alpha! = \alpha_1!\alpha_2 \ldots !\alpha_n!, \quad (x-a)^\alpha = h^\alpha = h_1^{\alpha_1} h_2^{\alpha_2} \ldots h_n^{\alpha_n}.$$

With this notation, the Taylor polynomial of order k has the formula:

$$T_{a,k}(h) = \sum_{\alpha:|\alpha|\leq k} \frac{h^\alpha \partial^\alpha f(a)}{\alpha!} \tag{6}$$

$|\alpha| = \alpha_1 + \ldots + \alpha_n$ is the order of the multi-index α. Thus the formula involves all derivatives of order up to k, including the value at the point, when $\alpha = (0, \ldots, 0)$.

The problem with the *Taylor* theorem is that the approximation error is not uniformly distributed. The approximation is accurate for x near a but becomes mediocre for larger values of x. But that won't be of concern since we'll be normalizing the features' variables all along the process, which will guarantee for us an optimal polynomial approximation.

In conclusion, the hypothesis space $\mathcal{H}$ that we'll be searching is the space of polynomials $\mathbb{P}[X]$ which contains the polynomial approximations of functions using Taylor's Polynomials.

3.2. *The reformulated supervised learning problem*

We assume that the model function we want to infer belongs to a polynomial space $\mathbb{P}[X]$ of degree of N. Therefore, problem (1) becomes:

$$\min_{\theta\in\mathbb{R}^N} \mathcal{J}(h) := \frac{1}{n}\sum_{i=1}^{n} max\left(1 - y_i\left(\sum_{k=0}^{N}\theta_k x_i^k\right), 0\right) + \lambda\|\theta\|^2. \tag{7}$$

In general, the problem (7) is ill-posed in Hadamard sense [28]. Consequently, the minimization stated in (1) is basically the same as resorting to the most common regularization method, the *Tikhonov* method [29].

The regularization of ill-posed problems, due initially to *Tikhonov* [27], seeks to redefine the notions of inversion and solution (quasi-solution, approximate solution,...) to make it, depends continuously on the data and is close to the exact solution (supposing that this exists for data close to the values actually obtained by the measurement). Concretely, the initial problem that is ill-posed is replaced by another, " close " to the first, and well-posed.

Solving an inverse problem using the regularization method of *Tikhonov* then comes down to a minimization problem, which requires the introduction of optimization techniques, specifically the descent gradient method. However, the optimization problem (7) is nonsmooth which does not allow the use of algorithms based on the gradient. To circumvent this problem the hinge loss function can be accurately approximated by a smooth function which is twice differentiable and convex [32]. For all $s \in \mathbb{R}$, the max function is approximated by a smooth function,

$$\max(s, 0) \simeq \ell_\gamma(s) := \left(s + \frac{\log(1 + \exp(-\gamma s))}{\gamma}\right), \quad \gamma > 0. \tag{8}$$

According to Lemma 1 in [32], we have that as γ tends to $+\infty$, the smoothed hinge loss ℓ_γ approaches the hinge loss ℓ.

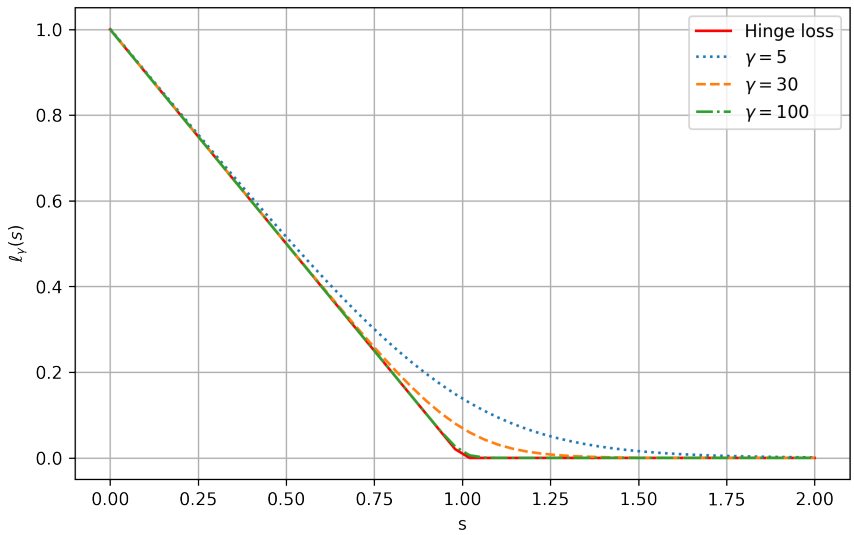


Figure 1. Smoothed hings loss

Then, ℓ_γ is used here to replace the the hinge loss ℓ in the problem (7). The approximate optimization problem of (7) is given by

$$\min_{\theta\in\mathbb{R}^N} \mathcal{J}_\gamma(h) \tag{9}$$

where

$$\mathcal{J}_\gamma(h) := \frac{1}{n}\sum_{i=1}^{n}\left(1 - y_i\sum_{k=0}^{N}\theta_k x_i^k + \frac{\log\left(1+\exp(-\gamma(1-y_i\sum_{k=0}^{N}\theta_k x_i^k))\right)}{\gamma}\right) + \lambda\|\theta\|^2.$$

Having approximate the nonsmooth problem (7) by smoothing (9), we will use the method of conjugate gradients at optimal pitch (base of the method of Fletcher-Reeves [30] to calculate the deformation parameter β_{k+1}) as a minimization tool. Before this, we need to compute the gradient of $\mathcal{J}_\gamma$ related to θ,

$$\nabla J_\gamma(h) = \left(\partial_{\theta_0}\mathcal{J}_\gamma, \ldots, \partial_{\theta_N}\mathcal{J}_\gamma\right)^T,$$

where $\partial_{\theta_k}\mathcal{J}_\gamma$ is given by

$$\partial_{\theta_k}\mathcal{J}_\gamma(h) = \frac{1}{n}\sum_{i=1}^{n} y_i x_i^k + \frac{y_i x_i^k \exp(\gamma(1-y_i\sum_{k=1}^{N}\theta_k x_i^k))}{1+\exp(\gamma(1-y_i\sum_{k=1}^{N}\theta_k x_i^k))}.$$

3.3. Nonlinear conjugate gradient method

Many nonlinear optimization methods can be used to solve the problem (9). The gradient-type methods are very attractive to solve this kind of optimization problem. The conjugate gradient method is one of the very effective gradient-type methods due to its simplicity and low storage [23, 26]. We briefly present the algorithm:

Algorithm 1: GC Method

1. For $m = 0,\ \theta^0 =$ given, compute $g_1 = \nabla\mathcal{J}_\gamma(h_{\theta^0})$ and $d_0 = -g_0$
2. while $\|\nabla\mathcal{J}_\gamma(h_{\theta^m})\| \geq \varepsilon$ do
   - (i) $m = m + 1$
   - (2i) $\alpha_m = arg\min_{\alpha\geq 0} J_\gamma(\theta^m + \alpha d_m)$;
   - (3i) $\theta^{m+1} = \theta^m + \alpha_m d_m$
   - (4i) $g_{m+1} = \nabla\mathcal{J}_\gamma(h_{\theta^m})$
   - (5i) $\beta_m = \| g_{m+1} \|^2 / \| g_m \|^2$
   - (6i) $d_{m+1} = -g_{m+1} + \beta_m d_m$
3. end.

In the following section we will move on to treat the experimental results obtained after the simulation of water potability data using the proposed approach.

4. Experimental results

To evaluate this approach it is often desirable to have some standardized benchmark data sets. In our case we choose to evaluate the proposed method through real-life data from from Kaggle called water potability [33]. To evaluate the efficiency of the proposed algorithm, we compared its accuracy with the K-Nearest Neighbors (KNN)[37, 38], Na?ve Bayes (NB) [34, 35] and Support Vector Machine (SVM)[31, 32] algorithms . The goal is to show that the proposed approach is faithfully to predict the considered class and we investigate on the importance of the features to find out which features contribute to most towards achieving the result.

The water potability data set used to train the models contained water quality metrics for 3276 different water bodies, and 10 parameter columns. The results of the research on the potability parameters are presented in the fig 2. Based on the results of observations regarding the percentage of drinking water quality based on the Water Potability database, from 3276 data that there is 59.67% water not consumed and 40.33% water can be consumed.

Attributes	description	Type of data
pH	pH of water	Numerical
Hardness	Capacity of water to precipitate soap in mg/L	Numerical
Solids	Total dissolved solids in ppm	Numerical
Chloramines	Amount of Chloramines in ppm	
Sulfate	Amount of Sulfates dissolved in mg/L	Numerical
Conductivity	Electrical conductivity of water in $\mu S/cm$	
Organic carbon	Amount of organic carbon in ppm	Numerical
Trihalomethane	Amount of Trihalomethanes in $\mu g/L$	
Turbidity	Measure of light emiting property of water in NTU (Nephelometric Turbidity Units)	Numerical
Potability	Indicates if water is safe for human consumption where 1 means Potable and 0 means Not potable	Numerical

Table 1. Water potability data set description.

Where ppm: parts per million, μg/L: microgram per litre and mg/L: milligram per litre.

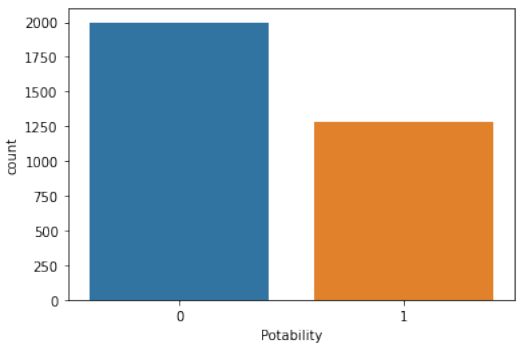


Figure 2. Data set class

The first 9 columns have been normalized to have mean 0 and Euclidean norm 1. For the implementation, we construct our training set by taking 2621 random observations form all data sets. Then by using the proposed approach, we generate predicting model.

Let's first find the correlation of each of these feature pairs and visualize the correlations using a heatmap. From Figure 4, there are three parameters that have a low relationship between parameters including Sulfate-Solids of 0.16 points and Sulfate-Hardness of 0.11 points. In addition, there are several parameters that do not have a close correlation between parameters. From the color indicator, we can read that the brighter the color, the greater the correlation between the two parameters, and vice versa if the color indicator is getting darker, the correlation between the two parameters is getting smaller. As seen from the figures 4-3 below, their is no linear relationship between the features as the plots are kind of circle. We can say that their is no multicollinearity.

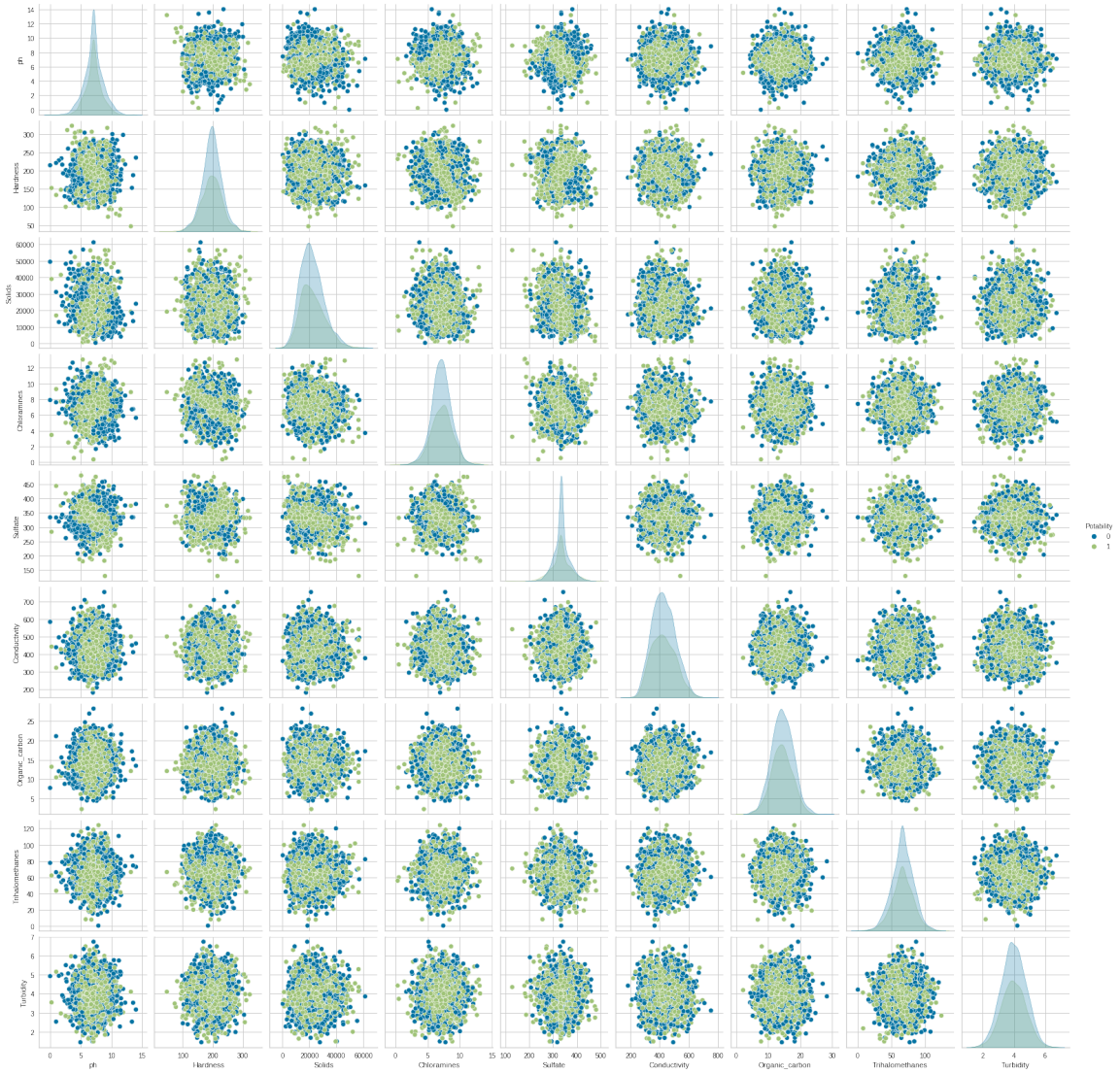


Figure 3. water potability dataset features.

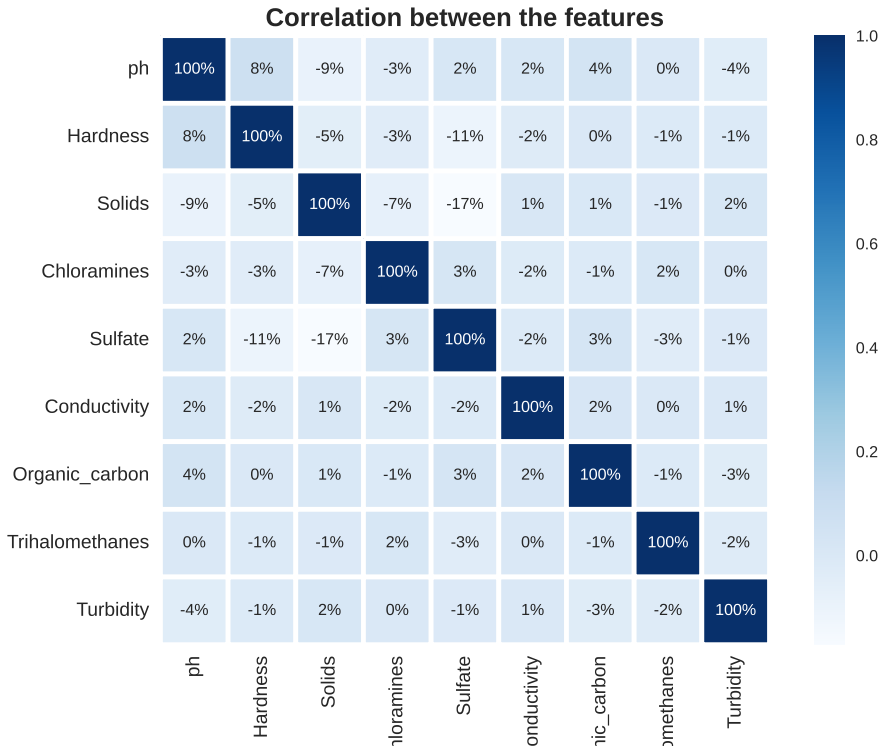


Figure 4. Correlation between the features in the water potability dataset.

In the sequel, for the proposed approach, we will use the polynomial approximation, given by

$$h(x) = \sum_{k=0}^{N} \theta_k x^k,$$

where $\theta_k \in \mathbb{R}$ and N is the degree of the polynomial.

Also, for the parameter settings, the degree of the polynomial N is set to be 4, while the parameter γ is adjusted to be 100, and the regularization parameter λ is chosen to be 10^{-3}. Furthermore, during the experiment, we used 80% of the data for training and 20% of the data for testing. All data points are normalized to $[0, 1]$ according to the minimum and maximum values in the corresponding feature column.

4.1. Results and discussions

The performance of the classification algorithms (KNN, SVM, DT and our method) was assessed in terms of the misclassification rate, sensitivity, specificity and accuracy of prediction, computed as below:

The accuracy of the test is its ability to Indicate if water is safe for human consumption or not cases correctly.

$$\frac{TP + TN}{TP + TN + FP + FN}.$$

The sensitivity (or Recall) of a test is its ability to Indicate if water is potable cases correctly.

$$\frac{TP}{TP + FN}.$$

The $F1$ **score** is a number between 0 and 1 and is the harmonic mean of precision and recall.

$$2 \times \frac{sensitivity \times precision}{precision + sensitivity}.$$

The precision of a test is its ability to determine the number of water samples the classifier labelled as potable that are actually potable

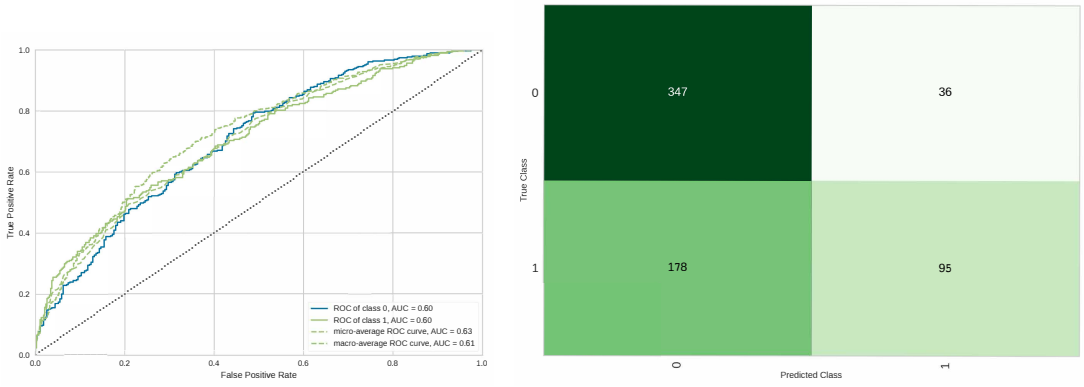
$$\frac{TP}{TP + FP}.$$

Where true Positive (TP) = the number of cases correctly identified as potable water, true negative (TN) = the number of cases correctly identified as not potable water, talse positive (FP) = the number of cases incorrectly identified as potable water and false negative (FN) = the number of cases incorrectly identified as not potable water.

Table 2 present the results of the used algorithms to predict water potability. It is noted that the performance of the proposed algorithm is very superior as compared to the SVM, KNN and NB. These results, allow us to say that the proposed algorithm has become more efficient in indicating if water is safe for human consumption or not.

	Accuracy	Sensitivity	Precision	$F1$ score
Our approach	**0.6737**	**0.6609**	**0.9016**	**0.7643**
SVM	0.5671	0.1654	0.4412	0.2406
KNN	0.6216	0.6087	0.4926	0.4009
NB	0.6324	0.5765	0.5292	0.25248

Table 2. Performance of the used algorithm to predict water potability.

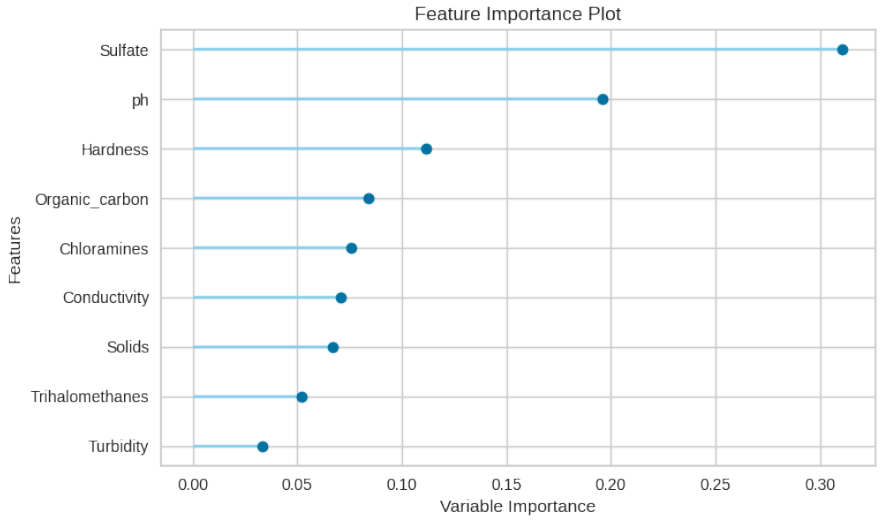


(a) (b)

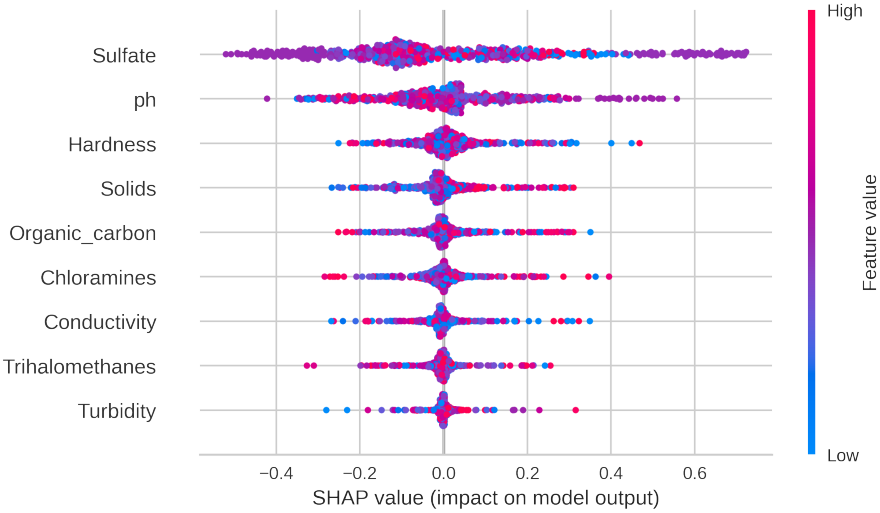
Figure 5. a) Roc Curves for the proposed approach. b) Confusion matrix for the proposed approach.

4.2. Feature importance

In this section, we investigate on the features importance to find out the factors which play important role in the water quality.



(a)



(b)

Figure 6. Feature importance.

From Figure 6, we see that, Sulfate is the most important feature, followed by pH.

5. Conclusion

In this paper, Tikhonov regularization and the Conjugate Gradient algorithm are proposed to predict Water Quality and therefore its potability. The comparison with the Support Vector Machine (SVM), the Naive Bayes (NB) and the K-Nearset Neighbours (KNN) algorithms on the basis of the accuracy and the precision shows the performance of the proposed algorithm in the prediction of the water potability and whether it is safe for human consumption or not. As a perspective we will focus on the integration of other methods for solving the optimization problem by exploring multiple loss functions as smoothing.

REFERENCES

1. Rochmi, MN. "Access to clean water is still very far from the target". Accessed from: https://beritagar.id/article/editorial/hapuskan-perda-causing-economy-high-cost. 2016.
2. UNESCO. "Global Climate Change". Accessed from: www.unesco.org. August 19, 2021.
3. Elysia, V. "Water and Sanitation where is Indonesia's position". National Seminar on the Role of Mathematics, Science, and Technology in achieving sustainable development goals/SDGs, FMIPA Open University. 2018. Pages 157-159.
4. Farooqui, A. "Invertigation of a community outbreak of typhoid fever associated with drinking water". BMC Public Health. 2009. 9:476.
5. Abrischamchi, A., Ebrahimian, A., Tajrishi, M., and Marino, M. "Case Study: Application of Multicriteria Decision Making to Urban Water Supply". J Water Resources Plan Manage. 2005. 131(4) Pages 326-335.
6. Cronin, A.A. "Monitoring source and domestic water quality in parallel with sanitary risk identification in Northern Mozambique to prioritise protection interventions". J Water Health. 2006. Volume 4 Pages 333-345.
7. Shyamala, R. "Physicochemical Analysis of Borwell Water Samples of Telungu Palayam Area in Coimvatore District, Tamilnadu, India". E-Journal of Chemistry. 2008. Volume 5(4), Pages 924-929.
8. Momba, M.N.M. "Abundance of pathogenic eschericia coli, Salmonella typhimurium and vibrio cholerae in Nkonkobe drinking water source". J Water Health. 2006. Pages 289-296.
9. Eschol, J. "Is fecal contamination of drinking water after collection associated with household water handling and hygiene practices? A study of urban slum households in Hyderabad, India". Journal of Water and Health. 2009. Pages 145-154.
10. C. Veesommai and Y. Kiyoki. "Critical Contaminate Detection, Classification of Multiple-water-quality-parameters Values and Real-time Notification by rSPA Processes". Surabaya ; IEEE International Electronics Sysposium, 2015.
11. A. Sarkar and P. Pandey, "River Water Quality Modeling Using Artificial Neural Network Technique". Aquatic Procedia, Vol.4, 2015, Pages 1070-1077.
12. Y. R. Ding, Y. J. Cai, P. D. Sun and B. Chen. "The Use of Combined Neural Networks and Genetic Algorithms for Prediction of River Water Quality" Journal of Applied Research and Technology. Vol.12, Issue 3, June 2014, Pages 493-499.
13. Noori, Roohollah, Zhiqiang Deng, Amin Kiaghadi, and Fatemeh Torabi Kachoosangi. "How Reliable Are ANN, ANFIS, and SVM Techniques for Predicting Longitudinal Dispersion Coefficient in Natural Rivers?" Journal of Hydraulic Engineering, 2015: 04015039.
14. Shuxiu Liang, Songlin Han, Zhaochen Sun, "Parameter optimization method for the water quality dynamic model based on data-driven theory". China, 2011.
15. Yue Liao, Jianyu Xu, Wenjing Wang, "A method of Water Quality Assessment Based on Biomonitoring and Multiclass Support Vector Machine". International Conferences on ESIAT, 2011.
16. Cortes C, Vapnik V."Support-vector networks". Machine learning. 1995;20(3):273C97.
17. Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, et al."Top 10 algorithms in data mining". Knowledge and information systems. 2008;14(1):1C37.
18. McCulloch WS, Pitts W. "A logical calculus of the ideas immanent in nervous activity". The bulletin of mathematical biophysics. 1943;5(4):115C33.
19. Webb G.I. (2011) Na?ve Bayes. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA.
20. Rosasco L., De Vito E., Caponnetto A., Piana M., Verri A. Are Loss Functions All the Same? Neural Computation. 16 (5), 1063-1076 (2004).
21. Lyaqini S., Quafafou M., Nachaoui M., Chakib A. Supervised learning as an inverse problem based on non-smooth loss function. Knowledge and Information Systems. 62, 3039-3058 (2020).
22. Lyaqini S., Nachaoui M., Quafafou M. Non-smooth classification model based on new smoothing technique. Journal of Physics: Conference Series. 1743, 012025 (2021).
23. S Lyaqini, M Nachaoui, A Hadri "An efficient primal-dual method for solving non-smooth machine learning problem" Chaos, Solitons & Fractals 155, 111754 (2022).
24. Nachaoui M. Parameter learning for combined first and second order total variation for image reconstruction. Advanced Mathematical Models and Applications. 5 (1), 53-69 (2020).
25. El Mourabit I., El Rhabi M., Hakim A., Laghrib A., Moreau E. A new denoising model for multi-frame super-resolution image reconstruction. Signal Processing. 132, 51-65 (2017).
26. Function minimization by conjugate gradients - Fletcher, Reeves - 1964.
27. Tikhonov, A. N., Arsenin, V. Y. Solutions to ill-posed problems. Winston-Wiley, New York (1977).
28. Hadamard, Jacques. Lectures on Cauchy's problem in linear partial differential equations. Courier Corporation, 2003.
29. S. Smale and D. Zhou. Shannon sampling II : Connections to learning theory. preprint, 2004.
30. Fletcher, R., and Reeves, C. M. (1964). Function minimization by conjugate gradients. The computer journal, 7(2), 149-154.
31. Lyaqini, S., Nachaoui, M., Quafafou, M. (2021). Non-smooth classification model based on new smoothing technique. In Journal of Physics: Conference Series (Vol. 1743, No. 1, p. 012025). IOP Publishing.
32. Lyaqini, S., Nachaoui, M. (2021, February). Identification of Genuine from Fake Banknotes Using an Enhanced Machine Learning Approach. In International Conference on Numerical Analysis and Optimization Days (pp. 59-70). Springer, Cham.
33. https://www.kaggle.com/adityakadiwal/water-potability.
34. Webb, G. I., Keogh, E., Miikkulainen, R. (2010). Na?ve Bayes. Encyclopedia of machine learning, 15, 713-714.
35. Zhang, H., Jiang, L., Yu, L. (2021). Attribute and instance weighted naive Bayes. Pattern Recognition, 111, 107674.
36. Otchere, D. A., Ganat, T. O. A., Gholami, R., Ridha, S. (2021). Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. Journal of Petroleum Science and Engineering, 200, 108182.
37. Khorshid, S. F., Abdulazeez, A. M. (2021). breast cancer diagnosis based on k-nearest neighbors: A review. PalArch's Journal of Archaeology of Egypt/Egyptology, 18(4), 1927-1951.

38. Ayyad, S. M., Saleh, A. I., Labib, L. M. (2019). Gene expression cancer classification using modified K-Nearest Neighbors technique. Biosystems, 176, 41-51.