IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 November 2023, accepted 29 November 2023, date of publication 5 December 2023, date of current version 18 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3339564

RESEARCH ARTICLE

Prediction of Water Quality Using SoftMax-ELM Optimized Using Adaptive Crow-Search Algorithm

S. R. SANNASI CHAKRAVARTHY[1], N. BHARANIDHARAN[2], VINOTH KUMAR VENKATESAN[2], (Member, IEEE), MOHAMED ABBAS[3], HARIKUMAR RAJAGURU[1], T. R. MAHESH[4], (Senior Member, IEEE), AND KRISHNAMOORTHY VENKATESAN[5]

[1]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam 638401, India
[2]School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore 632014, India
[3]Electrical Engineering Department, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia
[4]Department of Computer Science and Engineering, JAIN (Deemed-to-be University), Bengaluru 562112, India
[5]Department of Mathematics, Units of Basic Sciences, Arba Minch University, Sawla Campus, Arba Minch 4400, Ethiopia

Corresponding author: Krishnamoorthy Venkatesan (krishnamoorthy.venkatesan@amu.edu.et)

This work was supported by the Deanship of Scientific Research at King Khalid University (KKU) through the Research Group Program under Grant R.G.P.2/283/44.

ABSTRACT Water is a predominant source in the survival and development of all human lives. On top of all, predicting water quality is a significant one since water is essential in regulating our human body. In recent days, the advent of machine learning techniques has been supporting a lot in water quality prediction. Accordingly, Adaptive Crow Search Optimized SoftMax-Extreme Learning Machine (AdCSO-sELM) is proposed to improve the ELM performance by making the flight length adaptively with respect to the iterations. Here, the research novelty lies in making the CSOA parameters as a dynamic one which in turn provides promising ELM performance. Finally, the proposed AdCSO-sELM provides a superior accuracy of 96.54% for classifying water potability using the Kaggle dataset.

INDEX TERMS Crow search algorithm, extreme learning machine, neural network, optimization, water quality.

I. INTRODUCTION

Out of all natural resources, water is considered as the ultimate one for the survival and development of all lives. It is well-known that 66 percent of the earth is composed of water but evidently, only one percentage of water is consumable, and the remaining water is not consumable since they are either saline or salt water [1]. It is important to note that water plays a significant role in a country's wealth and economy. This directly provides an impact on both the utilization and development of water bodies. The changes in the quality of water are meticulously associated with human activities, seasonal deviations, and climatic environments. The quality variations in river water are indicative of gradual changes, and uncertain and non-linear factors. This makes the process of accurate water quality prediction more difficult [2]. At the same time, predicting water quality is highly significant for managing and planning of water resources and its environment. Based on the predicted outcomes, the prediction of the water pollution problem can be determined priorly and so the effects of water pollution can be avoided earlier. Furthermore, the primary growth of a nation highly depends on the abundant supply of water and its quality to preserve the ecosystem [3].

The accessibility of a safer and adequate amount of drinking water plays a critical role in the basic healthcare of human lives. This is because the quality of drinking water has a substantial impact on a healthy life [4]. The proportion of

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Gaggero.

elements in potable water needs not to threaten human life, and it needs to maintain the appropriate proportion regularly. The good quality of water should have the following attributes.

- Lower turbidity.
- Devoid of harmful agents.
- Lack of saline.
- No substances with an unpleasant smell or taste.
- Lower levels of hazardous elements such as lead which provide unsafe or negative health effects either in the short or long term.

A. ELEMENTS IN WATER

The standards to be followed in maintaining water quality differ among nations. As a common standard, the WHO provided guidelines for maintaining the lower and upper bounds of various inorganic elements existing in drinking water [5]. Accordingly, the sample admissible amount of elements [5] present in drinking water is listed in Table 1.

TABLE 1. Sample elements in water.

Elements in Water	Admissible Quantity (µg/L)
Fluoride	1500
Boron	2400
Barium	10
Arsenic	10
Chromium	50
Organic Species: Xylene	500
Organic Species: Benzene	10
Organic Species: Nitrilotriacetic acid	200
Organic Species: Dichloromethane	20
Organic Species: Styrene	20
Organic Species: Edetic acid	600
Organic Species: Toluene	700
Organic Species: Pentachlorophenol	9
Organic Species: Ethylbenzene	300

B. PROBLEM STATEMENT

All humans need to access safer and enough water for drinking, domestic use, food management, and recreational activities. A nation's economic growth might be substantially increased by better availability of water and a better ecosystem. Sufficient water availability for the usage of personal and domestic purposes is the basic right of every human being irrespective of nation. On every year, several people around the Earth face health issues such as kidney failure due to water contamination [6]. Additionally, the issue of contaminated water leads to typhoid, cholera, diarrhea, and other health problems [5]. Unfortunately, classifying water quality using laboratory procedures is time-consuming and requires more resources. However, a lot of systems are emerged in recent days for classifying water quality but they are not promising. Hence, a shed of light on the design of an automated as well as promising classification framework is the need of the hour for classifying water quality with minimal effort.

II. RELATED WORKS

The very important resource for the existence and survival of lives is water. All human needs are linked directly with water availability for several purposes such as drinking, catering, and other domestic usage [7]. Hence, it is highly significant in classifying water quality with state-of-the-art techniques. Several researchers have proposed their work toward the adopted research problem. The following discussion provides a review of previous research that employed artificial intelligence techniques for classifying water quality.

Juna et al. [8] proposed an automated classification of water quality based on KNN imputation and Multi-Layer Perceptron (MLP). The work efficiently took care of the missing attributes and yielded better classification performance. In the work, they used a 9-layered MLP model with KNN imputation. For comparison, the work utilized existing Machine Learning (ML) algorithms. The results suggested that their proposed model with KNN imputation provided an accuracy of 99% for predicting water quality. On the whole, the combination of the KNN imputer, MLP model, and thorough evaluation of different models contributes to the convincing results achieved by the proposed approach. Nasir et al. [9] developed a dependable approach for predicting water quality. The work utilized the existing ML algorithms such as decision tree, SVM, random forest, MLP, and CATBoost models. The work results that the CATBoost model yielded a superior accuracy of 94.5% for predicting water quality. Stacking ensemble models were applied, combining all classifiers, and so the above work provided convincing results. Another work proposed by Radhakrishnan and Pillai [10] used three ML models for predicting water quality. The algorithms used are Naïve Bayes, SVM, and decision tree and the work applied these algorithms on multiple datasets. Finally, they compared their performance and attained a better classification accuracy of 98.5% using the decision tree model. The efficient way of tuning parameters for the decision tree model contributes to the convincing results achieved by the proposed approach.

Aldhyani et al. [11] adopted long short-term memory (LSTM) and non-linear auto-regressive neural networks (NARNETs) models for predicting water quality. Additionally, the work employed KNN, NB, and SVM for experimental comparison. The results of LSTM and NARNET deep learning (DL) models overlapped with distinct regression results of RLSTM and NARNETs as 94.21% and 96.17%. Also, the SVM algorithm provided a maximum performance of 97% accuracy. According to the study, the kernel function employed for the SVM algorithm made the results more convincing than other models. For water distribution systems, a DL-based architecture was developed by Shahra et al. [12]. At this point, the work attained a higher accuracy with lower computation. Here, the research utilized SVM and ANN where the neural networks performed better than the SVM model, i.e., SVM and ANN yielded an accuracy of 89% and 94%. The superior results obtained for SVM in this research are due to the employment of the right choice of kernel functions.

Mohammed et al. [13] presented an adaptive neuro-fuzzy model for categorizing the inputs as safe or unsafe water for drinking purposes. The dataset used was a real-time based time-series comprised of pH, turbidity, color, and bacteria counts as attribute values. In this way, they got an accuracy result of 92% in detecting contamination inputs. The work utilized three membership function assignments for each input variable, resulting in convincing performance. Abuzir et al. [14] employed NB, MLP, and j48 ML models for classifying water quality. For this, they adopted a database comprised of ten distinct feature values. Additionally, the dimensionality reduction of the database was also done with different approaches. The work investigated 3 distinct cases: first is the consideration of all attributes, next is the consideration of only 5 feature values, and finally experimented with only two feature values. Using all feature values and carrying out feature selection, the work provided better performance for the MLP model. Hassan et al. [15] employed ML models for classifying the quality of Indian river water. The ML models include ANN, RF, SVM, multi-nomial log regression, and bagged-tree algorithms. The work exposed that the primary attribute values including nitrate, pH, dissolved oxygen, biological oxygen demand, and total coliform had a huge impact on water quality prediction. The feature correlation and selection to the employed machine learning classifiers contribute to the convincing results achieved by the proposed approach. The work of Sillbery et al. [5] utilized SVM and attribute realization (AR) models for predicting the water quality of the Chao Phraya River. While using the AR-SVM applied to 6 feature values of river water inputs, the research provided an accuracy range between 86-95%. This result is due to the integration of AR and SVM algorithms for the water quality prediction. The work of Ahmed et al. [16] utilized the dataset with four attributes of temperature, TDS, pH, and turbidity for predicting water quality. The outcomes of the experimentation revealed that the MLP model provided a better 85.05 % accuracy employing (3,7) configuration. A simpler MLP model with only four attributes contributes to the convincing results achieved by the proposed approach.

The systems employing IoT technology are highly used in the data acquisition of water quality databases. Kakkar et al. [17] employed the above technology for database creation in residential overhead containers. Once the creation of the database was done, the work utilized ML and DL models for predicting water quality. Due to the employment of a stacked model of ensemble H_2O AutoML algorithm, this work provided better results. Malek et al. [18] considered the data range between 2005 and 2020 in the Kelantan River Basin located in the Malaysian region. Here, they experimented with 13 chemical and physical attributes of water quality. As a final point, the experimentation revealed that the use of a gradient-boosting model with a 0.1 learning rate made the work to provide a maximum accuracy of 94.9%. Rustam et al. [19] employed a simple ANN architecture to predict water consumption as well as water quality. The work was investigated using two databases; one for predicting water quality and another one for water demand. The experimentation revealed that the ANN model yielded an accuracy of 96% whereas the R2 value was obtained as 0.99 for predicting water consumption. ANN with one hidden layer and a couple of dropouts and activation layers contribute to the convincing results achieved by the proposed approach.

A. RESEARCH OBJECTIVES

The research in water quality classification is an emergent one and the research community is working toward the best solution in order to support human lives. A simple water distribution system will not provide good health for human lives if the distributed water is not properly treated. And there may be a chance of contamination in the distribution system. Because of this contamination problem in drinking water, quality prediction has become a hot topic of interest for the survival of human lives over the past decade. Subsequently, a larger work toward predicting water quality has emerged in the literature. Comparatively, the research problem still lacks of having promising methodologies and so several methodologies have been proposed in recent days. Moreover, the available datasets for this research area have more missing values and this will lead to non-promising results. The proposed work intends to solve the above-said challenges using the subsequent contributions. At first, the cleaning of the dataset, i.e., the missing values is taken care of using a KNN imputer. By making use of this approach, the classification accuracy is significantly enhanced as compared to the research without imputation. Next, a novel hybrid classification architecture, Adaptive Crow Search Optimized SoftMax-Extreme Learning Machine (AdCSO-sELM) is proposed for the water quality classification. Here, the proposed method integrates the classification performance of ELM with the faster convergence of crow-search optimization. Furthermore, the crow-search optimization parameters are made as dynamic ones for attaining promising results.

Thus, the research work is organized in the following ways: this section describes studies related to the research problem. Section III describes the input dataset, visualization, and its preprocessing. The mathematical background of conventional ELM and Crow-Search optimization algorithms is discussed in Section IV. The proposed AdCSO-sELM is detailed in Section V whereas Section VI deliberates the discussion of results. Finally, the work concludes in Section VII.

III. MATERIALS AND METHODS

The discussion on the input dataset, data visualization, pre-processing, and proposed framework to classify water potability is done in this section. Figure 1 portrays the work-flow of the proposed framework used for predicting water quality. The work starts with the collection of a database that holds the acquired records of different attributes of water. The data are visualized and several missing values are found so pre-processed for tackling the missing values using KNN imputation. Then the proposed AdCSO-sELM architecture is applied to the research problem. Before this, the data is divided into testing and training sets. Finally, the performance of the proposed framework is analysed and compared with the existing approaches used for determining the quality of water as either potable or non-potable for humans.

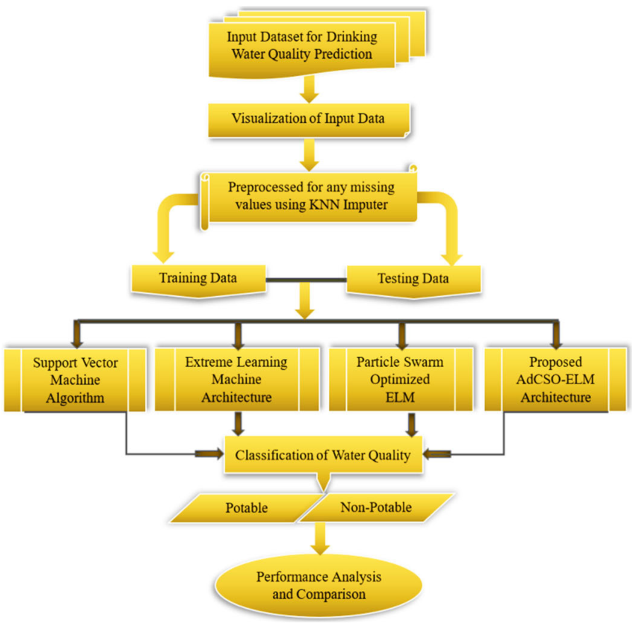

FIGURE 1. Work-flow for predicting water quality with AdCSO-sELM.

A. INPUT DATASET

The research employed data inputs taken from the Kaggle website [20]. This repository site is prominent for research data and is publicly accessible across the research community. The database utilized in this research is available on the above website as 'Water Quality'. The database comprises of ten columns with 935 instances of attributes. The output classes of the dataset are represented as either 1 or 0 where 0 indicates that the features representing the elements of water are not safer to drink and 1 depicts that the features illustrating the elements of water are safer to drink. This can be shortly represented as potable (1 as target) and non-potable (0 as target). The illustration of the above dataset is tabulated in Table 2.

TABLE 2. Attributes of input dataset with its description.

Attributes	Attribute Nature	Description of the Attributes
pH	Input (Independent)	Measurement of the potential of Hydrogen value ranges from 0→14.
Hardness		Indicator of soap precipitating ability in the water (measured in mg/L).
Solids		Overall dissolved solid elements measured in ppm.
Chloramines		Chloramine quantity (measured in ppm).
Sulfate		Sulfate quantity dissolved in water (measured in ppm).
Conductivity		Water's electrical conductivity (measured in µS/cm).
Organic Carbon		Organic carbon measurement in ppm.
Trihalomethanes		Trihalomethane measurement in µg/L.
Turbidity		Measurement of water clarity and quality.
Potability	Target (Dependent)	Value of '1' as potable and value of '0' as non-potable.

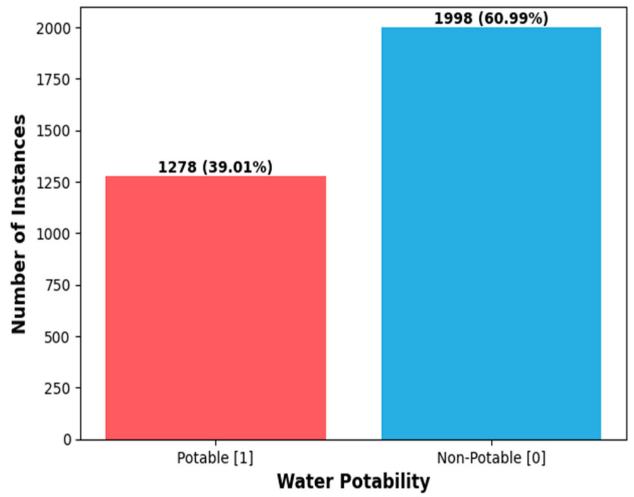

FIGURE 2. Distribution of output class targets of input dataset.

B. DATA VISUALIZATION

The visualization of the dataset supports scholars to expose the hidden relationships and patterns among the feature values. Figure 2 illustrates the data distribution of the employed water quality dataset. As in the plot, the number of features representing potability and non-potability is 1278 and 1998. Thus, the dataset contains 3276 instances with a binary class percentage of 60.99% and 39.01%. This plot reveals that the employed dataset is imbalanced so a robust classification model is required for the research problem. The next plot, Figure 3 illustrates the cluster map [21] plot of the input dataset. This plot provides a visual insight into the correlation among the feature inputs. As in the plot, the color scale denotes the correlation's direction and strength whereas the neutral color portrays a weaker correlation, the light and darker colors are a sign of positive and negative correlations. The obtained correlation coefficients are annotated over the colors. In this way, the plot reveals that the correlation between the target class and feature vectors has no direct correlation. This indicates that it is better to go ahead with all

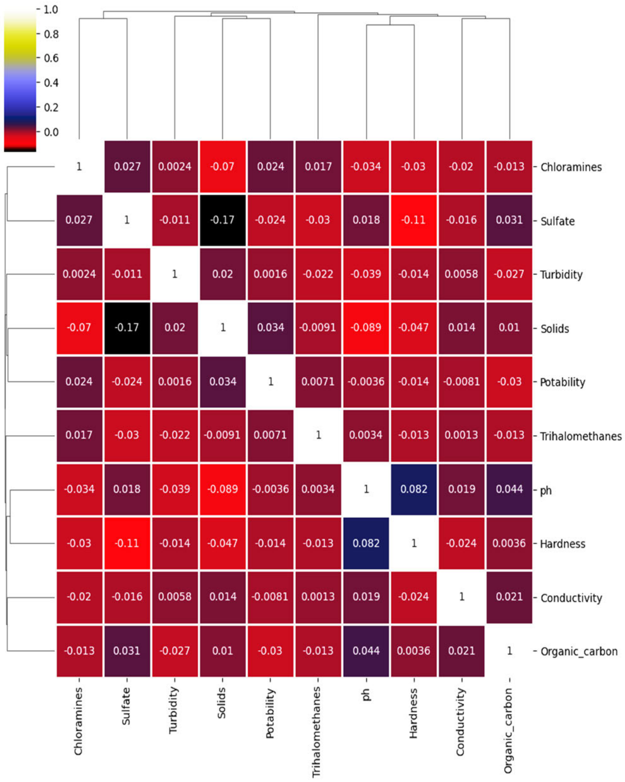

FIGURE 3. Cluster map visualization to illustrate correlation among data.

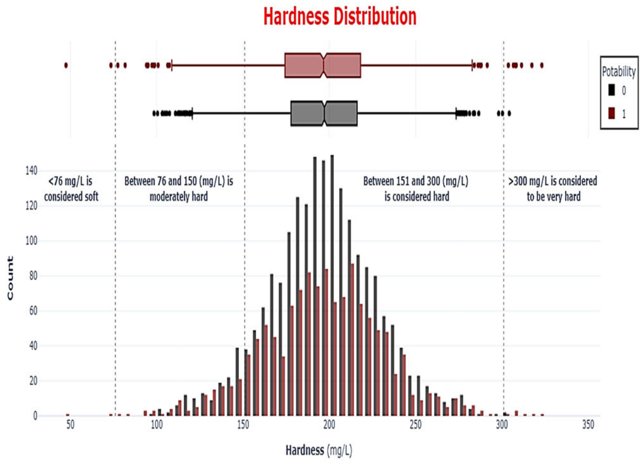

FIGURE 4. A sample characteristic visualization of hardness attribute of the dataset.

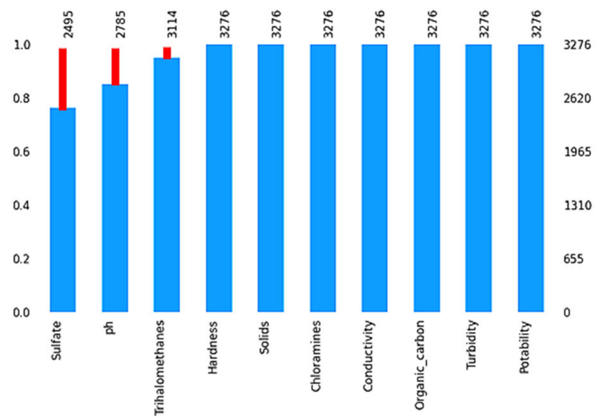

FIGURE 5. Bar plot for revealing the presence of missing values in input dataset.

the feature attributes for further classification stages. However, the absenteeism of correlation will lead to a challenge in performing the classification process due to higher non-linearity. Thus, a robust classifier is always a demand for the considered research work.

A sample characteristic visualization of the hardness attribute of the dataset is plotted in Figure 4. The plot creates a histogram showing the distribution of water hardness from the dataset. The histogram illustrates the distribution of water hardness with the corresponding frequencies. The x-axis represents water hardness in milligrams per liter (mg/L), while the y-axis represents the number of samples. The bars in the histogram are colored according to the drinkability of the samples. In addition to the histogram, three vertical lines are added at hardness levels of 76, 151, and 301 mg/L. The lines are shown as dashed lines and are used to classify the water hardness into different categories. Four annotations are also added to describe the different degrees of hardness.

The next plot of Figure 5 denotes a bar plot used to reveal the presence of missing values of the input dataset. As in the plot, the downloaded water quality dataset has missing values in three attributes out of nine which are sorted graphically in Figure 5. The attributes having missing values are sulfate (781), pH (491), and Trihalomethanes (162). The missing values are denoted as red bars. Thus, the framework requires a promising imputation technique for handling these missing values.

C. CLEANING OF MISSING VALUES USING KNN IMPUTER

Nowadays, data acquisition is done through several sources and utilized for insight analysis, validation, classification, and other processes. It is a more common problem that there may be a chance of missing information in the dataset since it deals with a huge volume of data. This, in turn, the usage of the created dataset to be limited due to missing values in one or more attributes. This might be due to several reasons such as the aging of components but the primary reason is human error. At this point, it is important to take care of null values in the dataset since they are important for prediction.

The total number of instances in the input dataset is 3276 as shown in Figures 2, 4, and 5. If the research does not take care of the missing values as represented in Figure 5, then the final prediction might be inaccurate since the influence of missed 1434 values changes the prediction to a different extent. Hence, the imputation of missing values is mandatory for any research problem. The study utilized the standard as well as popular KNN imputer [22] from the sci-kit library for handling the missing values of Figure 5. The technique

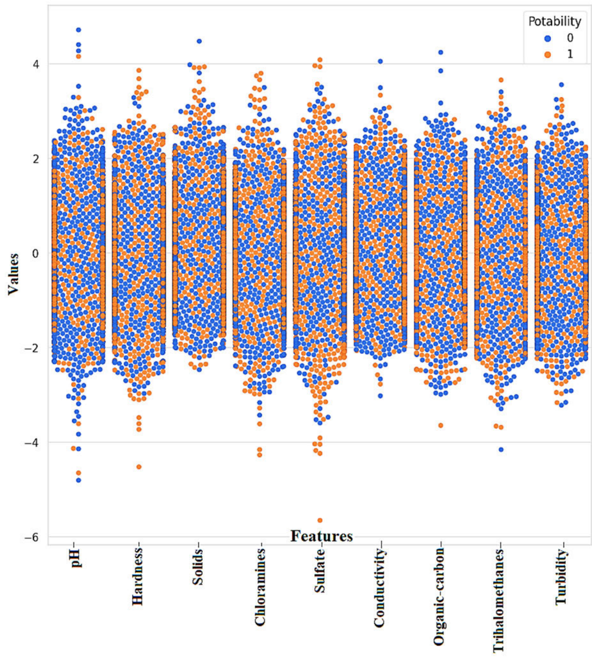

FIGURE 6. Swarm plot visualization of dataset after KNN imputation.

is widely used as an alternative to traditional imputation approaches. The idea behind the KNN imputation is to make use of Euclidean distance values for determining the nearest neighbors to replace missing values. Herein, the Euclidean distance is computed by increasing the weight of non-missed coordinates and not considering the missing values. The computation of Euclidean distance is mathematically defined as:

$$D_{xy} = \sqrt{Squared\ distance\ from\ present\ coordinates\ \times\ weight} \tag{1}$$

where

$$weight = \frac{total\ number\ of\ coordinates}{number\ of\ present\ coordinates} \tag{2}$$

The missing values are now imputed using the KNN algorithm and thus the dataset is complete for further visualization. The swarm plot [23] visualization of the dataset after KNN imputation is illustrated in Figure 6. This type of plot allows researchers to make data visualizations that plot numeric distributions of inputs against output targets effectively. From this seaborn plot, it is revealed that the input feature vectors employed are highly non-linear and overlap between the classes. Thus, the research employs robust ELM for further investigation of water potability.

IV. BACKGROUND OF ELM ARCHITECTURE AND CROW-SEARCH OPTIMIZATION ALGORITHM

A. EXTREME LEARNING MACHINE (ELM) ALGORITHM

The ELM architecture resembled the architecture of neural networks as given in Figure 7 but it is faster than the conventional neural networks [24]. For the employed classification problem, the computation of the hidden layer output matrix is mathematically given by [24]:

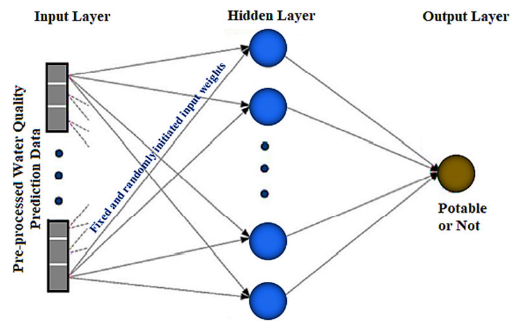

FIGURE 7. The architecture of ELM.

$$\boldsymbol{H\beta} = \boldsymbol{T} \tag{3}$$

where $\boldsymbol{T} = [y_1, y_2, \ldots y_L]^T$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots \beta_L]^T$ denote the output and weight vectors of the architecture. The hidden layer computation matrix can be expanded as [24],

$$\boldsymbol{H}\,(w_1 \ldots w_L, b_1 \ldots b_L, x_1 \ldots x_N) = \begin{bmatrix} \boldsymbol{G(w_1x_1+b_1)} & \cdots & \boldsymbol{G(w_Lx_1+b_L)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{G(w_1x_N+b_1)} & \cdots & \boldsymbol{G(w_Lx_N+b_L)} \end{bmatrix} \tag{4}$$

where $\boldsymbol{G}(.)$ is the ELM activation function and the final output weight solution is determined as:

$$\hat{\beta} = \boldsymbol{H}^{\dagger}\boldsymbol{T} \tag{5}$$

As in equation (5), $\boldsymbol{H}^{\dagger}$ denotes the Moore Penrose pseudo-inverse of $\boldsymbol{H}$. The research utilized the numpy library's *pinv function* to determine the $\boldsymbol{H}^{\dagger}$. Herein, the *pinv* function computes the generalized inverse of a matrix by making use of the singular-value decomposition technique (SVD). Finally, the function of the ELM algorithm is summarized as [25]:

(i) The input weights and hidden biases ($\boldsymbol{b_i}$ *and* $\boldsymbol{w_i}$) are randomly initialized.

(ii) Compute the values of the hidden layer output matrix, $\boldsymbol{H}$

(iii) Predict the output weights as $\boldsymbol{H}^{\dagger}\boldsymbol{T}$

(iv) Finally, $\hat{\beta}$ is taken to make a prediction on new inputs $\boldsymbol{T} = \boldsymbol{H}\hat{\beta}$

B. NEED FOR ELM OPTIMIZATION

The discussion in the previous sub-section reveals that the Extreme Learning Machine model is faster and more competent in providing solutions for several non-linear classification problems. At the same time, the characteristics of poor generalization and stability made ELM to yield poor classification performance. The randomly selecting nature of ELM's input weight and hidden bias parameters is the primary reason behind its reduced performance. Consequently, optimization of ELM is a need of the hour for getting better performances. Thus, a simple as well as adaptive crow-search optimization with faster convergence capability is adopted for optimizing ELM parameters namely input weights and hidden biases.

C. CONVENTIONAL CROW-SEARCH OPTIMIZATION ALGORITHM (CSOA)

It is important to note that special care is given in selecting an algorithm for effectively optimizing ELM parameters. If not, then the complexity of the overall framework might be increased with reduced performance. Thus, a simpler crow-search optimization algorithm involving the least parameter tunning (only two parameters) is employed.

In 2016, a researcher named Askarzadeh [26] introduced this algorithm which is influenced by the behavior of intelligent crows and it considers the following principles:

(i) All crows are to be involved as an individual of the group.

(ii) Any crow in the group can recall the concealed food spot.

(iii) Every crow will follow others in the group for stealing food.

(iv) However, all crows can probably remember and protect their food against others.

As discussed above, consider M as the size of the flock within a D-dimensional environment. And let the j^{th} crow's position with respect to the iteration as $y^{j,t}$ where j ranges from 1 to M. From the above principles of algorithm, every crow should effortlessly remember the hidden food spots; let $N^{j,t}$ indicates the hidden place memory of j^{th} crow at t iterations and t_{max} represents the maximum iterations. Also, $N^{j,t}$ denotes the so far best position attained by the j^{th} individual.

Consider that at t iteration, a j^{th} crow starts to track the obscured food-spot of z^{th} crow. This probably provides the following cases [26]:

Case 1: The z^{th} crow has no idea about j^{th} crow's tracing of its obscured food spot. Subsequently, the j^{th} crow approaches the z^{th} crow's obscured food. This makes the position of the j^{th} crow as an updated one at further $(t+1)$ iterations as [26],

$$y^{j,t} = y^{j,t+1} + R_j \times fl^{j,t} \times (N^{z,t} - y^{j,t}) \tag{6}$$

Here, R_j indicates a random number generated through uniform distribution [0, 1] and $fl^{j,t}$ denotes the j^{th} crow's flight length with respect to t iterations.

Case 2: In another scenario where the z^{th} crow has enough awareness about j^{th} crow's tracing of its obscured food spot. Because of this, the z^{th} crow intends to mislead j^{th} crow within the search space.

Finally, the above cases are mathematically represented as [26]:

$$y^{j,t+1} = \begin{cases} y^{j,t} + R_j * fl^{j,t}\left(N^{z,t} - y^{j,t}\right), R_z \geq AP^{j,t} \\ \text{else choose a random position} \end{cases} \tag{7}$$

Here, R_z indicates a random number generated through uniform distribution [0, 1] and $AP^{j,t}$ is the awareness probability. The value of AP decides the better convergence of the crow-search algorithm within the search space.

The optimization process starts with initializing the parameters: D, M, t_{max}, fl *and* AP respectively. The study considers the M value as the number of input feature vectors of the employed pre-processed dataset. Next, the value of y illustrated in equations (6) and (7) represents the position of each crow and is randomly initialized within the space. During execution, the crows' position will be evaluated and checked through the use of a fitness function (RMSE). Consequently, all crows within the search-space should start to update their positions as mentioned in equation (7). However, the position updation is examined for its feasibility. And its mathematical representation representing fitness function (Fn) can be given as [26]:

$$N^{j,t+1} = \begin{cases} N^{j,t+1}, & Fn\left(y^{j,t+1}\right) \text{ is better than } Fn\left(N^{j,t}\right) \\ N^{j,t}, & \text{otherwise} \end{cases} \tag{8}$$

V. THE PROPOSED METHOD: ADCSO-SELM ALGORITHM

The ELM model is optimally tuned using crow-search optimization having two parameters namely awareness probability (AP) and flight length (fl). Additionally, the proposed model is enhanced by making the above parameter (fl) as an adaptive one with respect to the iteration.

The impact of parameter, fl in updating the position of crows inside the search space is illustrated in Figure 8. Here, if the flight length value is lesser than 1 as in Figure 8(a), then the next position of j^{th} crow will be located in between $y^{j,t}$ and $N^{z,t}$. And if the flight length value is greater than 1, then the next position of j^{th} crow will be located outside the line as given in Figure 8(b). Thus, the algorithm's search ability highly depends on the value of fl. If fl is too larger, then the algorithm tends to search globally and has poor convergence. In the next case, if fl is too smaller, the algorithm could easily fall into local optima.

The aforesaid point reveals that the value of fl can be set critically to balance the tradeoff between exploration and exploitation. It is also noted that the crows have good memory capability by nature so that they can hide around thousands of foods at different locations during rainy or autumn seasons and they can fly differently to retrieve their food. And in the algorithm, the above point is valid only if the fl parameter is set as dynamic. Hence, for imitating the crow's natural intelligence, fl in the research problem is made as adaptive with respect to the iterations ($Adfl$).

The flowchart of the proposed algorithm is portrayed in Figure 9. As in Figure, the crow's natural intelligence is mimicked by making the fl as a time-varying parameter. As a result, in order to support random exploration in the whole search-space, the value of fl is made to be higher at initial iterations. Once initiated, the value of fl is made to be lower for increased iterations to improve the exploitation of a better solution attained so far. This is due to the chance

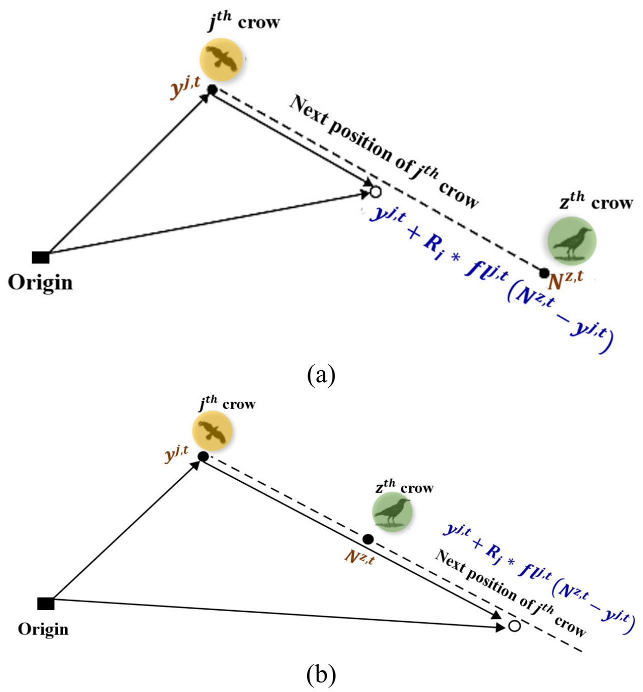

FIGURE 8. Position updation of j^{th} crow with (a) $fl < 1$ (b) $fl > 1$.

of attaining better solutions near the current solutions would become higher in the successive iterations. In this way, the *fl* parameter of the conventional CSA algorithm is made as adaptive during the execution of the algorithm. In short, the value of *Adfl* would start with an initial value and later become smaller progressively with time i.e., iterations. Thus, the mathematical representation can be expressed as:

$$Adfl = \left(\frac{Iter}{t_{max}} Adfl_{min}\right) + Adfl_{max} - \left(\frac{Iter}{t_{max}} Adfl_{max}\right) \quad (9)$$

Here, the minimum and maximum bounds of *fl* during the iteration are represented as $Adfl_{min} = 0.1$ and $Adfl_{max} = 2$. Additionally, *Iter* and t_{max} represent the current iteration and maximum iteration count (100). By using the above equation, a better balance of exploration and exploitation is attained in the work. Figure 10 illustrates how the flight length varies with respect to iterations where the crow's flight length is made adaptive between 0.1 and 1.981. Additionally, the plot reveals that the value of *fl* is decreasing as time progresses as discussed earlier. It is also to be noted that the adaptive flight length does not have a substantial impact on the complexity of the proposed framework. The reason behind this is that *fl* will be calculated for the solutions using equation (9) while updating each crow's position during every iteration of the algorithm.

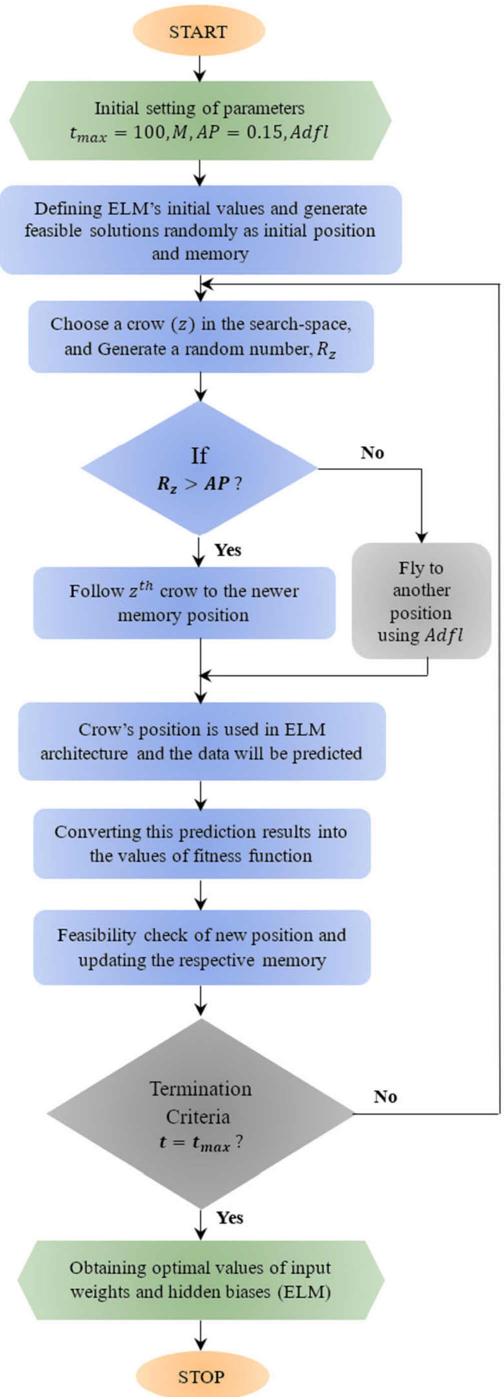

FIGURE 9. Flowchart of AdCSO algorithm.

VI. EXPERIMENTAL STUDY AND ANALYSIS

The outcomes of the classification of water quality using the proposed method are discussed here. The research execution is implemented using Python 3.6 IDE installed on a Windows 10 OS system with 8 GB RAM and Intel i7 processor. For comparative analysis, the study employed the four existing algorithms namely standalone ELM, Support Vector Machine (SVM) [27], Naïve Bayes (NB) [25], K-Nearest Neighbors (KNN) [27], Multilayer Perceptron (MLP)) [25], Particle Swarm Optimized ELM [28], and Conventional Crow-Search Optimized ELM algorithms. For the classification phase, the work utilized the stratified partition of a 70:30 ratio of training and testing set of inputs.

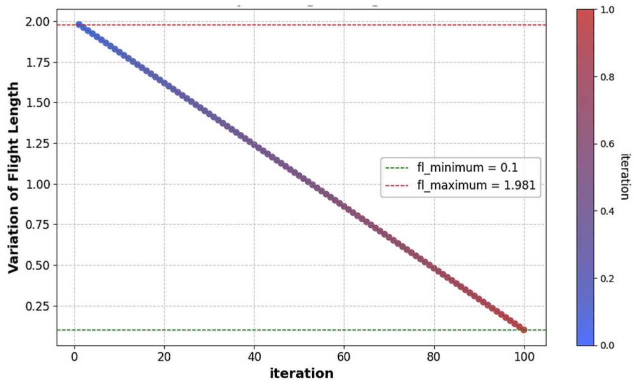

FIGURE 10. Time-varying flight length of CSA.

A. EXPERIMENTATION ON SELECTION OF ACTIVATION FUNCTION FOR ELM

The algorithmic implementation of ELM always has the challenge in its selecting the number of hidden neurons [29]. Additionally, for handling the employed non-linear dataset (as shown in Figure 6), the research needs to select the appropriate activation function. The research experiments seven distinct activation functions and hidden neurons that ranged up to 350. The experimentation is carried out using conventional ELM architecture. Figure 11 shows the ELM's mean accuracy (5-fold cross-validation) comparison for different activation functions and hidden neurons.

As in Figure 11, the experimentation of selecting the number of hidden neurons and activation functions [30] includes the employment of SoftMax, tanh, sigmoid, ReLU, leaky ReLU, parametric ReLU (pReLU), and ELU functions. Based on the experimental plot, the research selects the number of hidden neurons as 100 and the type of activation function as softmax. This is because the selected activation function provides superior performance at each selection of hidden neurons at both the training and testing phases. The SoftMax activation function is actually a soft version of argmax function and it is mathematically defined as [31]:

$$y_i = \frac{e^{x_i - \max(x)}}{\sum_{j=1}^{n} e^{x_j - \max(x)}} \tag{10}$$

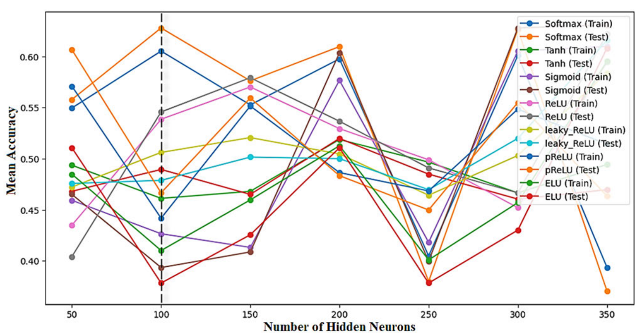

FIGURE 11. ELM's mean accuracy comparison for different activation functions and hidden neurons.

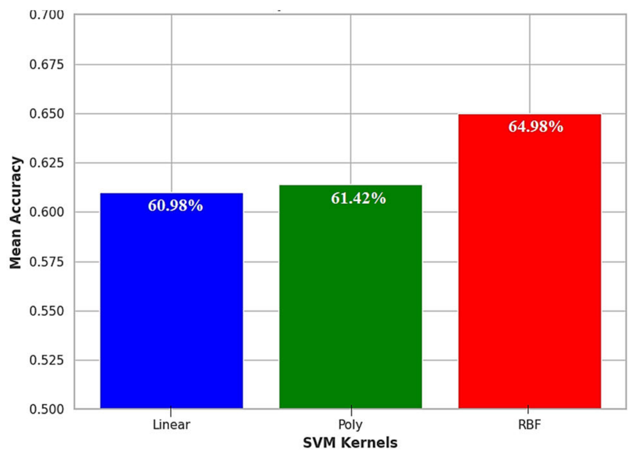

FIGURE 12. SVM's mean accuracy comparison of different kernel functions.

Here, the numerator is used for calculating the exponential of difference between the i^{th} element and the maximum element, $\max(x)$ in the input vector. The denominator is for calculating the sum of the exponentials of all elements in the input vector after adjusting for the maximum value. The idea of subtracting the $\max(x)$ from each element before exponentiating is for preventing potential issues with an overflow that could occur usually while exponentiating larger values. This calculation does not affect the final result because it cancels out in both the numerator and denominator of equation 10. Thus, the output of the above equation is a vector with probabilities of each feasible outcome.

B. RESULTS OF THE PROPOSED WORK AND ITS COMPARATIVE ANALYSIS

For classification problems, the literature study reveals that the Support Vector Machine is the most widely used algorithm next to Neural network algorithms. For handling non-linear inputs (Figure 6), the SVM with three distinct kernel functions namely linear, poly, and radial basis function (RBF) [32] is used. The comparison of mean accuracies obtained using 5-fold cross-validation is plotted in Figure 12. This plot reveals that the SVM employed with the RBF kernel function (Mean accuracy of 64.98%) provides superior performance over others.

For PSO-ELM, the control parameter values are set are as follows: cognitive factor (c_1), social factor (c_2), and interia weight (w) are typically tuned as 2, 2, and 0.7. Additionally, as discussed before, the SoftMax activation function provides a better non-linearity as compared with others. Thus, PSO-sELM is employed for the comparative performance analysis of the study. As given in Figure 9, the initial parameters of the proposed AdCSO-sELM are set for obtaining the experimental results. The study employed the performance metrics that are derived using the confusion matrix (CM). This CM contains the elements of four parameters namely true positive and negatives, false positive and negatives represented as TP, TN, FP, and FN. The employed metrics are classification accuracy, sensitivity, specificity, and precision. To attain a better balance between precision and sensitivity, the F1 score metric is employed for the analysis. And for further validation, Cohen's kappa metric [33] is adopted for validating the performance analysis.

TABLE 3. Classifier's performance for classification of water potability.

Classifiers	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)	F1 Score (%)
NB	66.44	60.57	64.15	72.50	69.34
KNN	66.61	60.84	64.36	72.68	69.51
MLP	67.28	62.14	65.27	73.54	70.27
SVM-RBF	67.11	61.62	64.98	73.22	70.03
sELM	66.78	59.27	63.85	71.94	69.26
PSO-sELM	81.14	75.20	78.82	83.65	82.37
CSO-sELM	88.15	86.42	87.47	91.03	89.57
Proposed AdCSO-sELM	95.83	97.65	96.54	98.46	97.12

The attained results for the employed binary classification problem after 5-fold cross-validation are tabulated in Table 3. Herein, the NB, KNN with K = 3 provided a classification accuracy of around 64.15-64.36%. But the performance of MLP neural network is higher than the above-mentioned classifiers with the performance of 67.28% of sensitivity, 62.14% of specificity, 65.27% of accuracy, 73.54% of precision, and 70.27% of F1 score. Also, noted that the performance of the SVM classifier employed with the RBF kernel and ELM algorithm employed with the SoftMax activation function provides an overlapping performance. That is, both these classifiers provide an accuracy of 63.85-64.98%, 71.94-73.22% of precision, 66.78-67.11% of sensitivity, 59.27-61.62% of specificity, and 69.26-70.03% of F1 score. However, the performance of traditional SVM-RBF is slightly better than the conventional sELM architecture in the water potability problem. This is due to the better generalization ability of SVM for unseen data and thus provides slightly better performance than ELM architecture. This makes the SVM classifier as a benchmark and popular choice for several classification problems.

For further improving the performance, the conventional sELM is optimized using particle swarm optimization (PSO-sELM), and due to this optimization, the performance of sELM has been elevated from 63.85% to 78.82% in terms of classification accuracy. In addition, the resultant values of precision and F1 score are elevated to 83.65% and 82.37%. This, in turn, confirms that ELM is faster but needs appropriate optimization for selecting input weights and hidden biases. Herein, PSO tends to focus more on exploration, which might lead to faster convergence but might miss out on fine-tuning the solution in complex search spaces. In general, PSO demands careful tuning of parameters, and in dynamic environments, PSO might struggle to adapt effectively. The adaptive nature of the CSOA parameters could make the optimization well-suited for dynamic and changing environments. This adaptability allows CSOA to adjust its flight lengths with respect to iterations, potentially leading to better convergence in our employed work. Thus, the research is moving forward to obtain superior performance using sELM together with better optimization. In this way, the simple and efficient crow-search algorithm is adopted for optimizing sELM (CSO-sELM).

The CSO-sELM provided a better classification performance than PSO-sELM, that is, the algorithm provided a sensitivity of 88.15%, specificity of 86.42%, accuracy of 87.47%, precision of 91.03% and F1 score of 89.57%. This is found to be the better performance obtained so far for the employed binary classification problem. This is because of the ability of crow-search optimization algorithm to explore a wider search space with the potential for global optima. However, the research problem always needs to attain the best performance. Hence, the optimization process of crow-search algorithm is further fine-tuned, i.e., one of the parameters, flight length is made to be adaptive throughout iterations as shown in Figure 10. During the experimentation of the proposed algorithm, here also, the SoftMax activation provided better non-linear mapping of features over other functions. Consequently, the proposed AdCSO-sELM provides a superior classification performance of sensitivity (95.83%), specificity (97.65%), accuracy (96.54%), precision (98.46%), and F1 score (97.12%).

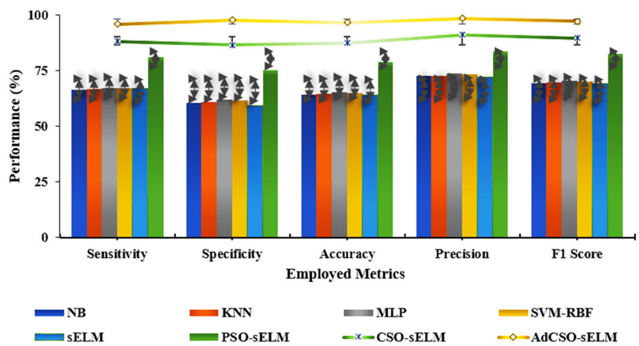

FIGURE 13. Performance comparison of proposed algorithm for water potability problem.

The above discussion is graphically compared in Figure 13 and reveals that the proposed AdCSO-sELM provides the algorithm-best performance over others in classifying water potability. This plot also includes the visualization of the standard deviation error bar for all classifiers. The standard deviation (SD) error bars of Figure 13 provide information on how spread the performances are around the mean value. Herein, the smaller bars represent the lower spread i.e., the performances are clumped around the mean whereas large bars represent the higher spread i.e., the performances are more variable from the mean. In this way, the conventional NB, KNN, MLP, SVM-RBF, and sELM have been reported with larger errors over other algorithms. The CSO-sELM has been reported with smaller SD errors whereas the proposed algorithm provides the least SD errors.

C. STATISTICAL KAPPA ANALYSIS OF THE PROPOSED RESEARCH

For further statistical validation, the study employs Cohen's kappa (κ) coefficient to validate the attained results. For the employed binary classification problem, κ coefficient provides a statistical evaluation of agreement between classifiers when classifying instances into either of two classes (potable or not potable). The κ supports in assessing the reliability of the classification process beyond what would be expected by chance alone. Herein, $\kappa = 0$ represents that the observed agreement is the same as would be expected by chance alone (least performance). The κ value ranged between 0.00 - 0.20 representing slight agreement, 0.21 - 0.40 representing fair agreement, 0.41 - 0.60 representing moderate agreement, 0.61 - 0.80 representing substantial agreement, and the value ranged between 0.81 - 1.00 represents almost perfect agreement.

The comparison of accuracy with kappa values is plotted in Figure 14. In this plot, the NB, KNN, MLP, SVM-RBF and sELM ($\kappa = 0.26\ to 0.29$) algorithms have a fair agreement in classifying water potability whereas the PSO-optimized sELM algorithm has a moderate agreement ($\kappa = 0.56$) in this binary classification task. The CSO-optimized sELM algorithm provides a substantial agreement ($\kappa = 0.74$) and the proposed AdCSO-sELM algorithm provides an almost perfect agreement ($\kappa = 0.93$) among others for predicting water quality. Thus, the results obtained are validated using κ metric that validates the superior performance of the proposed algorithm in classifying water quality as discussed in Table 3 and Figure 13.

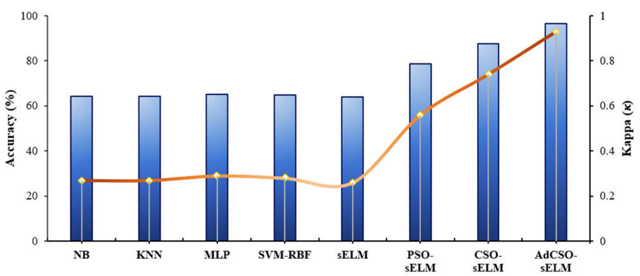

FIGURE 14. Accuracy and kappa comparison of the proposed method.

D. COMPARISON WITH STATE-OF-THE-ART FRAMEWORKS

Finally, the proposed methodology is compared against the existing research studies and it is summarized in Table 4. For this, the recent research works that are associated with the employed research problem are chosen. The listed research works utilized the distinct water quality or monitoring dataset for classifying water potability. As from the comparison findings summarized in Table 4, the proposed methodology outperforms other frameworks due to the right choice of classification model with appropriate parameter tuning for the water quality prediction task.

TABLE 4. Comparison of the proposed framework with the recent research works for predicting water quality.

Research Work	Year	Algorithm	Overall Classification Accuracy (%)
Aldhyani et al. [11]	2020	RLSTM	94.21
Hasan et al. [34]	2021	Decision Tree	95.00
Dilmi et al. [35]	2021	LDA with LSTM-RNN	88.00
Shahra et al. [12]	2021	ANN	94.00
Rustam et al. [19]	2022	LSTM	89.00
Afaq et al. [8]	2022	XGBoost	76.00
Hamza et al. [36]	2023	Gradient Boosting Machine	80.00
Proposed Method	2023	AdCSO-sELM	96.54

E. EVALUATION USING FURTHER DATASETS

To further evaluate the robustness of the proposed framework, the research makes use of two more larger and complex datasets. These datasets named WQ_1 [37] and WQ_2 [38] are taken from the Kaggle web repository. The WQ_1 dataset comprises of 21 columns out of which the last one is the output target. This deals with checking the quality of water as either safer or not for human activities with 7,996 instances. This includes the feature attributes that are the composition of elements in water such as uranium, silver, selenium, radium, perchlorate, mercury, nitrites, nitrates, lead, virus, bacteria, fluoride, copper, chromium, chloramine, cadmium, barium, arsenic, ammonia, and aluminium components. The next dataset, WQ_2 consists of 10,48,575 instances with the elements of Air Temperature, Water Temperature, Total Dissolved Solids, Manganese, Chlorine, Conductivity, Sulfate, Odor, Copper, Fluoride, Turbidity, Color, Zinc, Lead, Chloride, Nitrate, Iron, and pH values. This dataset also deals with checking the quality of water as either safer or not for human activities. In addition to this information, the abovementioned two datasets have the missing values. And these missing values are imputed using the KNN algorithm as discussed in Section III. In a similar way, the results obtained after applying the proposed AdCSO-sELM algorithm with a similar set-up are comparatively analysed in Figure 15.

From the plot of Figure 15, it is evident that the proposed AdCSO-sELM algorithm provides an improved classification performance as compared with other models. That is, the proposed framework provides an improved accuracy of 97.37% for the WQ_1 dataset and 95.73% for the WQ_2 dataset, respectively. This performance is appropriately validated using Kappa statistical analysis where improved kappa values of 0.87 and 0.89 are obtained for the above datasets.

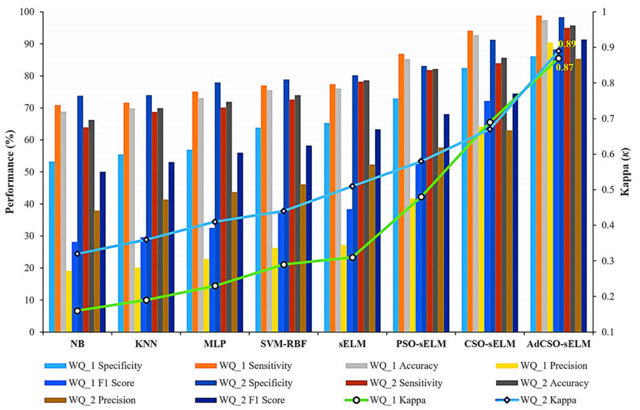

FIGURE 15. Comparative analysis of the proposed framework with WQ_1 and WQ_2 datasets.

Thus, the proposed model works well with any newer case of inputs employed for water quality prediction.

F. DISCUSSION OF FINDINGS

- The summary of the study intended to enhance the ELM performance for the water potability classification problem is given below.
- The research novelty lies in enhancing the classification performance of sELM through the successful implementation of the proposed Adaptive Crow-Search Optimized SoftMax-Extreme Learning Machine (AdCSO-sELM).
- The work utilized the Crow-search algorithm as compared with the firefly algorithm [39], grey wolf optimizer [40], dragonfly algorithm [41], grasshopper optimization algorithm [42], Harris Hawks optimization [43], Hybrid Capuchin and Rat swarm algorithm [44] due to its adaptive nature and comparatively easier to implement or tune for the employed requirements of the water quality prediction task.
- The above framework is then utilized to improve the s-ELM's classification performance in predicting the quality of water as safer (potable) or not (non-potable).
- As compared with recent existing methods, the proposed algorithm performed well for the employed binary classification task.
- The study compared the classification performance of the proposed framework with conventional NB, KNN, MLP, ssSVM-RBF, s-ELM, PSO-sELM, CSO-sELM and reported that the AdCSO-sELM outperforms them and thereby establishing the novelty of the framework.

G. LIMITATIONS OF THE PROPOSED FRAMEWORK

Several real-world societal problems always require more precise results using machine learning or deep learning algorithms. In this regard, the research community is working towards providing the best solutions. In this way, the proposed framework is successfully implemented and promising results have been obtained. However, as from the plot of Figure 13, the sensitivity of the proposed framework is reported to be slightly lower than the specificity. This reveals that the ability of the proposed methodology needs to be improved to correctly classify unsafe (non-potable) water samples minimizing the risk of classifying unsafe water as safe. Consequently, this problem should be taken care of in our future research.

VII. CONCLUSION AND FUTURE WORK

Nowadays, the world is running towards more sophistication and this has a severe impact on the environment. This contamination provides a negative effect on human health and so becomes a cause of several diseases. Due to this, finding appropriate water for drinking purposes is essential for human life. The study in the research intended to design a promising framework for predicting water as either potable or not for drinking purposes. For evaluation, the work utilized the Kaggle water quality dataset. The missing values in the dataset are imputed using the KNN algorithm. After visualization, the dataset is found to be highly non-linear, and subsequently, the SoftMax function is chosen as an activation function of ELM architecture. Based on the experimentations, the performance of ELM can be optimally improved using a simple meta-heuristic crow-search algorithm. This, in turn, the research proposed an Adaptive Crow-Search Optimized – SoftMax Extreme Learning Machine (AdCSO-sELM) for obtaining robust performance for the employed binary classification problem. Accordingly, the performance of AdCSO-sELM is compared with the existing frameworks and found to be more competent than others. That is, the proposed framework provides a superior classification accuracy of 96.54% and a precision of 98.46% with a kappa validation of 0.93. Additionally, the proposed framework is evaluated using two more larger and complex datasets and found that the AdCSO-sELM model is robust for water quality prediction. The future work will be in the extension of evaluating the proposed framework with real-world time-series data in an IoT environment. Furthermore, it is intended to employ multi-class data for further evaluation with deep-learning architectures.

FUNDING

This research received no external funding.

DATA AVAILABILITY STATEMENT

The dataset used for the findings is included in the manuscript.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest exists.

REFERENCES

[1] T. Yan, S.-L. Shen, and A. Zhou, "Indices and models of surface water quality assessment: Review and perspectives," *Environ. Pollut.*, vol. 308, Sep. 2022, Art. no. 119611.

[2] S. R. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "A systematic review on machine learning algorithms used for forecasting lake-water level fluctuations," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 24, p. e7231, Nov. 2022.

[3] S. Giri, "Water quality prospective in twenty first century: Status of water quality in major river basins, contemporary strategies and impediments: A review," *Environ. Pollut.*, vol. 271, Feb. 2021, Art. no. 116332.

[4] R. Bain, R. Johnston, and T. Slaymaker, "Drinking water quality and the SDGs," *Npj Clean Water*, vol. 3, no. 1, p. 37, Aug. 2020.

[5] C. Sillberg, P. Kullavanijaya, and O. Chavalparit, "Water quality classification by integration of attribute-realization and support vector machine for the Chao Phraya River," *J. Ecological Eng.*, vol. 22, no. 9, pp. 70–86, Oct. 2021.

[6] R. A. Kristanti, T. Hadibarata, M. Syafrudin, M. Yilmaz, and S. Abdullah, "Microbiological contaminants in drinking water: Current status and challenges," *Water, Air, Soil Pollut.*, vol. 233, no. 8, p. 299, Aug. 2022.

[7] Z. Luo, Q. Shao, Q. Zuo, and Y. Cui, "Impact of land use and urbanization on river water quality and ecology in a dam dominated basin," *J. Hydrol.*, vol. 584, May 2020, Art. no. 124655.

[8] A. Juna, M. Umer, S. Sadiq, H. Karamti, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "Water quality prediction using KNN imputer and multilayer perceptron," *Water*, vol. 14, no. 17, p. 2592, Aug. 2022.

[9] N. Nasir, A. Kansal, O. Alshaltone, F. Barneih, M. Sameer, A. Shanableh, and A. Al-Shamma'a, "Water quality classification using machine learning algorithms," *J. Water Process Eng.*, vol. 48, Aug. 2022, Art. no. 102920.

[10] N. Radhakrishnan and A. S. Pillai, "Comparison of water quality classification models using machine learning," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2020, pp. 1183–1188.

[11] T. H. H. Aldhyani, M. Al-Yaari, H. Alkahtani, and M. Maashi, "Water quality prediction using artificial intelligence algorithms," *Appl. Bionics Biomechanics*, vol. 2020, pp. 1–12, Dec. 2020.

[12] E. Q. Shahra, W. Wu, S. Basurra, and S. Rizou, "Deep learning for water quality classification in water distribution networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.* Cham, Switzerland: Springer, Jun. 2021, pp. 153–164.

[13] H. Mohammed, I. A. Hameed, and R. Seidu, "Machine learning: Based detection of water contamination in water distribution systems," in *Proc. Genetic Evol. Comput. Conf. Companion*, Jul. 2018, pp. 1664–1671.

[14] S. Y. Abuzir and Y. S. Abuzir, "Machine learning for water quality classification," *Water Quality Res. J.*, vol. 57, no. 3, pp. 152–164, Aug. 2022, doi: 10.2166/wqrj.2022.004.

[15] M. M. Hassan, M. M. Hassan, L. Akter, M. M. Rahman, S. Zaman, K. M. Hasib, N. Jahan, R. N. Smrity, J. Farhana, M. Raihan, and S. Mollick, "Efficient prediction of water quality index (WQI) using machine learning algorithms," *Hum.-Centric Intell. Syst.*, vol. 1, nos. nos. 3–4, pp. 86–97, 2021.

[16] U. Ahmed, R. Mumtaz, H. Anwar, A. A. Shah, R. Irfan, and J. García-Nieto, "Efficient water quality prediction using supervised machine learning," *Water*, vol. 11, no. 11, p. 2210, Oct. 2019.

[17] M. Kakkar, V. Gupta, J. Garg, and S. Dhiman, "Detection of water quality using machine learning and IoT," *Int. J. Eng. Res. Technol. (IJERT)*, vol. 10, no. 11, pp. 73–75, 2021.

[18] N. H. A. Malek, W. F. Wan Yaacob, S. A. Md Nasir, and N. Shaadan, "Prediction of water quality classification of the Kelantan River basin, Malaysia, using machine learning techniques," *Water*, vol. 14, no. 7, p. 1067, Mar. 2022.

[19] F. Rustam, A. Ishaq, S. T. Kokab, I. D. L. T. Diez, J. L. V. Mazón, C. L. Rodríguez, and I. Ashraf, "An artificial neural network model for water quality and water consumption prediction," *Water*, vol. 14, no. 21, p. 3359, Oct. 2022.

[20] Kaggle. (2021). *Water Quality*. Accessed: Jun. 22, 2022. [Online]. Available: https://www.kaggle.com/datasets/adityakadiwal/water-potability

[21] C. L. Gilchrist and Y. H. Chooi, "Clinker & clustermap.js: Automatic generation of gene cluster comparison figures," *Bioinformatics*, vol. 37, no. 16, pp. 2473–2475, 2021.

[22] P. Keerin and T. Boongoen, "Improved KNN imputation for missing values in gene expression data," *Comput., Mater. Continua*, vol. 70, no. 2, pp. 4009–4025, 2022.

[23] M. Burch, G. Wallner, N. Broeks, L. Piree, N. Boonstra, P. Vlaswinkel, S. Franken, and V. van Wijk, "The power of linked eye movement data visualizations," in *Proc. ACM Symp. Eye Tracking Res. Appl.*, May 2021, pp. 1–11.

[24] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[25] S. R. S. Chakravarthy and H. Rajaguru, "Detection and classification of microcalcification from digital mammograms with firefly algorithm, extreme learning machine and non-linear regression models: A comparison," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 1, pp. 126–146, Mar. 2020.

[26] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm," *Comput. Struct.*, vol. 169, pp. 1–12, Jun. 2016.

[27] A. Roy and S. Chakraborty, "Support vector machine in structural reliability analysis: A review," *Rel. Eng. Syst. Saf.*, vol. 233, May 2023, Art. no. 109126.

[28] S. R. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Processing of digital mammogram images using optimized ELM with deep transfer learning for breast cancer diagnosis," *Multimedia Tools Appl.*, vol. 82, no. 30, pp. 47585–47609, Dec. 2023.

[29] S. R. S. Chakravarthy and H. Rajaguru, "Automatic detection and classification of mammograms using improved extreme learning machine with deep learning," *IRBM*, vol. 43, no. 1, pp. 49–61, Feb. 2022.

[30] C. Abirami, R. Harikumar, and S. R. S. Chakravarthy, "Performance analysis and detection of micro calcification in digital mammograms using wavelet features," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2016, pp. 2327–2331.

[31] V. Vinoth Kumar, L. Wang, J. I.-Z. Chen, B. Sikdar, and M. Nones, "Guest editorial: Trends, perspectives and prospects of sensor technologies in hydrological sciences," *Acta Geophysica*, vol. 70, no. 6, pp. 2837–2839, Nov. 2022.

[32] G. N. Kouziokas, "SVM kernel based on particle swarm optimized vector and Bayesian optimized SVM in atmospheric particulate matter forecasting," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106410.

[33] G. M. Foody, "Explaining the unsuitability of the Kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification," *Remote Sens. Environ.*, vol. 239, Mar. 2020, Art. no. 111630.

[34] A. N. Hasan and K. M. Alhammadi, "Quality monitoring of abu dhabi drinking water using machine learning classifiers," in *Proc. 14th Int. Conf. Develop. eSystems Eng. (DeSE)*, Dec. 2021, pp. 1–6.

[35] S. Dilmi and M. Ladjal, "A novel approach for water quality classification based on the integration of deep learning and feature extraction techniques," *Chemometric Intell. Lab. Syst.*, vol. 214, Jul. 2021, Art. no. 104329.

[36] H. A. Madni, M. Umer, A. Ishaq, N. Abuzinadah, O. Saidani, S. Alsubai, M. Hamdi, and I. Ashraf, "Water-quality prediction based on H_2O AutoML and explainable AI techniques," *Water*, vol. 15, no. 3, p. 475, Jan. 2023.

[37] Kaggle. (2021). *Water Quality Dataset for Water Quality Classification*. Accessed: Nov. 9, 2023. [Online]. Available: https://www.kaggle.com/datasets/mssmartypants/water-quality/data

[38] Kaggle. (2021). *Water Quality Prediction*. Accessed: Nov. 9, 2023. [Online]. Available: https://www.kaggle.com/datasets/vanthanadevi08/water-quality-prediction/

[39] Z. Zhou, J. Wang, Z. Zhu, and J. Xia, "Robot manipulator visual servoing based on image moments and improved firefly optimization algorithm-based extreme learning machine," *ISA Trans.*, early access, 2023, doi: 10.1016/j.isatra.2023.10.010.

[40] Z. Zhang and Z. Zhou, "Illumination correction via improved grey wolf optimizer for regularized random vector functional link network," *KSII Trans. Internet Inf. Syst.*, vol. 17, no. 3, Mar. 2023, doi: 10.3837/tiis.2023.03.008.

[41] J. Li, W. Shi, and D. Yang, "Clothing image classification with a dragonfly algorithm optimised online sequential extreme learning machine," *Fibres Textiles Eastern Eur.*, vol. 29, no. 3, pp. 91–96, Jun. 2021.

[42] J. Li, W. Shi, and D. Yang, "Color difference classification of dyed fabrics via a kernel extreme learning machine based on an improved grasshopper optimization algorithm," *Color Res. Appl.*, vol. 46, no. 2, pp. 388–401, Apr. 2021.

[43] J. Li, W. Shi, and D. Yang, "Fabric wrinkle evaluation model with regularized extreme learning machine based on improved Harris hawks optimization," *J. Textile Inst.*, vol. 113, no. 2, pp. 199–211, Feb. 2022.

[44] M. A. Baseer, V. V. Kumar, I. Izonin, I. Dronyuk, A. K. Velmurugan, and B. Swapna, "Novel hybrid optimization techniques to enhance reliability from reverse osmosis desalination process," *Energies*, vol. 16, no. 2, p. 713, Jan. 2023.

S. R. SANNASI CHAKRAVARTHY received the B.E. degree in ECE from P. T. Lee CNCET, Kanchipuram, India, in 2010, the M.E. degree in applied electronics from the Thanthai Periyar Institute of Technology, Vellore, India, in 2012, and the Ph.D. degree in the domain of machine learning from Anna University, Chennai, in 2020. Currently, he is an Associate Professor with the Bannari Amman Institute of Technology, India. His research interests include machine learning and deep learning techniques for solving various real-world problems.

N. BHARANIDHARAN received the Ph.D. degree from the Faculty of Information and Communication Engineering, Anna University, Chennai, in 2020. He is currently an Assistant Professor with the School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, India. His research interests include machine learning and deep learning techniques for solving various real-world problems.

VINOTH KUMAR VENKATESAN (Member, IEEE) is currently an Associate Professor with the School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, India. He is the author/coauthor of papers in international journals and conferences, including SCI-indexed papers. He has published over 60 articles in IEEE Access, Springer, Elsevier, IGI Global, and Emerald. His current research interests include wireless networks, the Internet of Things, machine learning, and big data applications. He is an Associate Editor of *International Journal of e-Collaboration* (IJeC) and *International Journal of Pervasive Computing and Communications* (IJPCC) and an Editorial Member of various journals.

MOHAMED ABBAS received the B.Sc. degree in electronics engineering from the Faculty of Engineering, Mansoura University, Egypt, in 1998, the M.Sc. degree in computer engineering from Mansoura University, in 2002, and the Ph.D. degree in computer engineering from Mansoura University, in 2008. Since this date, he has been worked as an Assistant Professor with the Department of Communications and Computer Engineering, College of Engineering, Delta University. He is currently working as a Full Professor with the Department of Electrical Engineering, King Khaled University, Abha, KSA. His areas of interest are intelligent systems, medical informatics, nanotechnology, and bioinformatics.

HARIKUMAR RAJAGURU received the Ph.D. degree in the domain of bio-signal processing from Anna University. He is currently a Professor with the Bannari Amman Institute of Technology, Sathyamangalam, India. His research interests include machine learning and deep learning for solving various real-world problems.

T. R. MAHESH (Senior Member, IEEE) is currently an Associate Professor and the Program Head of the Department of Computer Science and Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, India. He has to his credit more than 40 research articles in Scopus and SCIE-indexed journals of high repute. He has been an editor for books on emerging and new-age technologies with publishers, such as Springer, IGI Global, and Wiley. His research interests include image processing, machine learning, deep learning, artificial intelligence, the IoT, and data science. He has served as a reviewer and a technical committee member for multiple conferences and journals of high reputation.

KRISHNAMOORTHY VENKATESAN received the bachelor's and master's degrees in mathematics discipline from the University of Madras, Tamil Nadu, India, the M.Phil. degree in mathematics from Thiruvalluvar University, and the Ph.D. degree in mathematics (fuzzy algebra) discipline from Bharathiar University, Tamil Nadu, in 2016. In addition, he was qualified for the eligible test of lectureship examination at the state level, in 2017. He is currently an Associate Professor with Arba Minch University, Sawla Campus, Ethiopia. His research interests include fuzzy and intuitionistic fuzzy sets and its related fields.