International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

Analysis of Support Vector Machine Model in Forecasting the One Year Ahead Water Quality of the River Ganga

Anil Kumar Bisht[1], Ravendra Singh[2], Rakesh Bhutiani[3], Ashutosh Bhatt[4]

[1,2]Computer Science & Information Technology, Mahatma Jyotibha Phule Rohilkhand University, Bareilly, U.P. India
[3]Deptt. Of Environment Science, Gurukul Kangri University, Haridwar, Uttarakhand, India
[4]Deptt. Of Computer Science, BIAS, Bhimtal,Uttarakhand, India

Abstract: Analysis of river water quality forecasting model based on Support Vector Machine (SVM) for the Ganges River has not been reported in the literature to the best knowledge of the authors. This paper attempts to develop such models for the Ganges River in the area from Devprayag to Roorkee, Uttarakhand, India. A monthly experimental dataset has been used from 2001 to 2015 time series. This paper examines the possibilities of SVM techniques for the 2016 water quality forecast of the Ganges River based on a continuous dataset of the last fifteen years. In conclusion, the SVM-based technique failed to develop a year ahead forecasting models (as compared to Artificial Neural Network (ANN) based previous work) that could efficiently estimate the water quality of the Ganges River. In the future, some other modifications will be made in the proposed model in order to have efficient forecasting.

Keywords: Support Vector Machine (SVM), Mean Square Error (MSE), Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD) and Water Quality (WQ).

1. INTRODUCTION

Water is the basic need for any nation to live. Since it is a fundamental part of what is required for both to live as natural life everywhere around the world. In any case, the nature of water is quickly being uprooted due to the chaotic industrialization, augmented population, substantial and unstructured urbanization and so on, as well as man-made issues of late. As we know, rivers are major water wells. The Ganges River is the national river of India but in recent years its water quality is also deteriorating day by day. Therefore, it is the need of the hour to become aware of such a dangerous situation and to take technically sound steps to bring the river Ganga to its sacred condition. [1].

Estimating the water quality of rivers is an extraordinary fundamental miracle in itself and is also a difficult undertaking. The advancement of such competent determination models is an exceptionally dull task due to the complex and non-relationship between factors responsible for deciding water quality. Accurate expectation of water quality would indicate the advancement of the system and options as a whole by water management authorities. During the last few decades it has been a tremendous development in the field of artificial intelligence (AI), specifically targeting the problem of pattern recognition. AI is one of such a growing area of SVM that has been successfully applied in forecasting or forecasting problems. Based on historical data, SVMs are able to perform non-linear mapping between a given data set and its elements i.e. identifying trends and patterns in the given data to produce the desired output. ANNs technique for water quality forecasting has been analyzed by the authors [1]. SVM is analyzed in this paper on the same dataset for the Ganges River. The purpose of this paper is to analyze and develop various forecasting models in one year advance for the Ganga River water quality based on the SVM approach using a time series experimental dataset for a period of fifteen years from 2001 to 2015.The remaining segments of the paper are organized in five sections. Related work is discussed in section 2, the adopted methodology is described in next section 3. The section 4 consists of results and analysis while the entire work is concluded in the section 5.

2. RELATED WORK

Recently, many researchers and scientists applied methods of ANN and SVM in their respective domains that included modeling and prediction problems. Some of them are the Istanbul Stock Exchange Price Index Predictions based on ANN and SVM [2], whether predicted using NN and SVM [3], intrusion detection in combination with ANN technique with SNM method [4], an assortment of merlot wines based on SVM and ANN [5], ANN and SVM [6] predict the metabolizable energy required in the compound feed for pigs. In addition to these applications of ANN and SVM models, these concepts have also been applied in the area of water quality (WQ) modeling of rivers, like the application of ANN and SVM methods in predicting dissolved oxygen [7], [8]. Author in [9] applied SVM for bio-monitoring and water quality assessment based on a multi-class support vector machine. The experiments were conducted on behavior of fish based on computer vision techniques and multi-classification method. Authors concluded that their results were found to be satisfactory.

Thus, the study of the literature led to the conclusion that despite the importance of ANN and SVM-based models in various fields, including WQ prediction, no such technique is still applicable for modeling Indian rivers and forecasting their water quality. In relation to this, not a single research has been done (to the best of knowledge) to include such technical solutions for the WQ of India's national river Ganga, whose water quality is deteriorating day by day. Keeping all these facts in mind, this chapter deals with the development of the WQ-forecasting model for the Ganges River, which incorporates the techniques of SVM, which are implemented in the Weka data mining tool.

3. RESEARCH METHODOLOGY

3.1 Study Region and the Dataset Used

The Ganges river basin in India has been taken up for research work in this paper. Monthly experimental data sets collected for the WQ of the Ganges River at five different stations in Uttarakhand: Devprayag, Rishikesh, Haridwar, Jwalapur and Roorkee analyzed by the Limnology and Ecological Modeling Laboratory in Uttarakhand, Gurukul Kangri University, Haridwar, Uttarakhand. This data set comprised of various WQ parameters. Temperature, pH, dissolved oxygen (DO) and biochemical oxygen demand (BOD) are taken for the experiments. The Ganges river basin is as shown in figure 1.

Fig. 1 The Ganga River Basin (Source: Google Earth)

The current research work aims to analyze the performance of the water quality forecasting model developed for the Ganges River based on SVM technique on historical and time-series datasets collected with five sampling sites over a period of fifteen years from 2001 to 2015. The Weka tool [10] has been used for experimentation. This is particularly a type of tool used in data mining for prediction classification, clustering. The details of the data set collected is given in the table 1. A snapshot of the dataset is displayed in figure 2. The dataset displayed is for Station A Devprayag for the year 2001. In the same style it has been collected in the same format for all five stations till the year 2015. Once the entire received dataset is systematically organized as tables in Excel format, independent and dependent variables are identified i.e. input factors (features) that affect the WQ and corresponding target variables of the Ganges River are responsible for.

Table 1 Water Quality Data for River Ganga, Uttarakhand, India

Year of Data Used: 2001 to 2015	
Frequency of Collection: Monthly	
Station Sr. No.	**Station Name**
1.	Devprayag
2.	Rishikesh
3.	Haridwar
4.	Jwalapur
5.	Roorkee

			2001				
	DATA	AT SAMPLING POINT-A (DEVPRAYAG)					
Sr. No.	Month/Year	Temp °C	pH	DO (mg/L)	BOD (mg/L)	Total Coliform /100ml	WQ
1	Jan-01	14.5	7.42	12.6	1.8	94	2
2	Feb-01	14.8	7.68	11.8	2	70	2
3	Mar-01	15.2	7.7	11	2.4	110	2
4	Apr-01	15	8.03	11.5	1.8	110	2
5	May-01	16.5	8.08	11.2	2	120	2
6	Jun-01	16.8	8.19	10.7	2.9	140	2
7	Jul-01	17.1	8.81	11	3.6	160	4
8	Aug-01	16.5	8.7	12.1	3	120	2
9	Sep-01	16.4	8.64	11.9	2.6	120	2
10	Oct-01	15.4	8.3	11.2	1.8	110	2
11	Nov-01	15.8	7.96	11	1.6	90	2
12	Dec-01	14.9	7.61	12	1.2	110	2

Fig. 2 A sample of data set

The dependent variable is also called the target variable that represents the output to be determined or tested and its value depends on the independent variable also known as the feature variable whose value can be manipulated and which in turn affects the result of the experiment. In the context of machine learning, the predictor variable refers to the input variable that needs to be mapped to the output or target variable [11]. Therefore, according to these definitions, four variables temperature, pH, dissolved oxygen (DO) and biochemical oxygen demand (BOD) are identified as independent and taken in this work, while one variable water quality is determined as dependent.

Temperature: This is an important factor affecting water quality. WQ factors such as dissolved oxygen and biochemical oxygen demand are greatly affected by changes in temperature. The lower temperature cools the water and therefore the level of DO in that water is higher and vice versa.
Dissolved oxygen (DO): It is a marker of the amount of oxygen present in water that sinks either through the atmosphere or through plant photosynthesis. DO levels must be maintained high in water as it is an important element for aquatic life.
Biochemical-Oxygen Demand (BOD): It represents the total DO required for the breakdown of organic waste or material by the organism. BOD levels mean that the amount of waste present in the water consumed by bacteria is high and thus pollution is also high.
pH: The pH level of water indicates the acidity or alkalinity of water. It actually shows the capacity of hydrogen i.e. measures the amount of hydrogen combined with water. A value of 7 is the balanced pH known as neutral, alkalinity increases below and above this acidity.

Water Quality: It defines the characteristics of water in relation to physical, chemical and biological considerations. There are various categories in which water quality is classified for its specific uses. In this work we have followed the water quality standards recommended by the "Central Pollution Control Board of India" (CPCB).

3.2 Process of Model Development using SVM in Weka

Based on the technique of SVM all the experiments are performed in the WEKA tool for developing and analyzing the performance of various WQ-forecasting models for the Ganges River.

Support Vector Machine

SVM provides a learning method aimed at pattern recognition. A classifier built on SVM divides the mentioned data sample into two distinct categories by sketching a straight line termed as a hyper-plane to implement the classification task.

Let us consider a set of input dataset known as input vector: [2].

$(a_1, b_1), (a_2, b_2), (a_3, b_3) \dots (a_n, a_n)$:

a_i belongs to R^d while i= 1,2,3,……n

which denotes vectors 'n' of dimension-'d' where the output dataset equivalent to this is termed as labels as: b_i belongs {-1,1}. The primary function of SVM involves determining a hyper-plane (assuming the feature vector A and label B) so that it makes the best classification in such a way that -1 and +1 of the input data with two output classes [9] Represents classification. The main focus of SVM is to draw the hyper-plane as a separate boundary with the goal that it maximizes the separation between the data points surrounding the two ranges with respect to this different plane. These data points that belong to two different categories are support vectors [12]. This distance is defined by a term margin as highlighted in the figure 3.

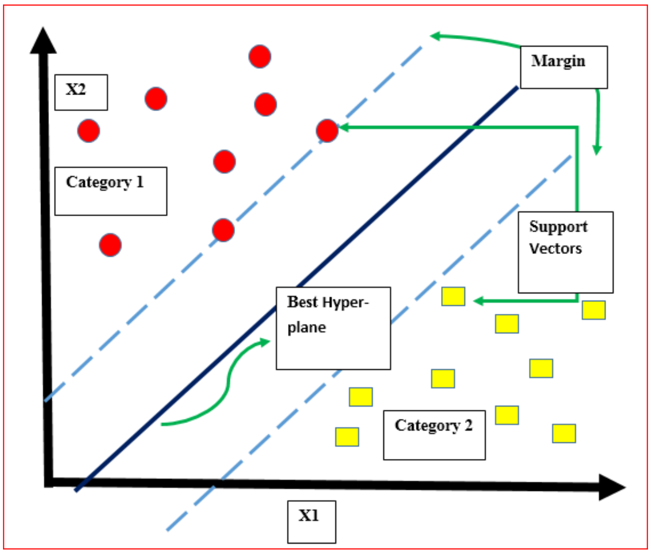

Fig. 3 Support Vector Machine [12]

All in all, the majority of this present reality issues are not linear in nature i.e. they can't be categorized into directly divisible categories as above. To handle such cases the nonlinear SVM based on kernel function approach is much helpful where corresponding to the input dataset its equivalent high-dimensional feature space is created by applying nonlinear transformation. Then, a normal ideal hyperplane is to be established that separates data into two categories. A kernel is defined as a mathematical function used in SVMs. Its main role is to implement the concept of non-linear SVM for supporting the issues in the real world involving non-linearity. The kernel function accept the input data and map it into the desired form. In general the kernel function involves transformation of input data into the other dimension that will provide a clear separating margin among the different classes of the input data. A number of kernel functions are there like Polynomial function, Radial-Basis Function (RBF), sigmoid function etc. Among these the most ordinarily utilized by the scientists is the RBF.

The RBF is defined as follows:

$K(a, a_i) = e^{-\gamma || a - a_i ||^2}$, where γ represent the constant in RBF whose value need to be larger than zero and $| a-a_i |$ represents the distance from x center to the x_i input vector as Euclidean distance [13].

After selecting the RBF kernel, a comma separated value (CSV) file of the dataset is created. Secondly, this file is converted into a Weka compatible file format as attribute-relation file format (arff). A supervised machine learning approach, called SVM, is used to create a predictive model by setting its parameters in Weka. Now in the next step the various files with the dataset in the arff file imported into Weka. Next, we configured the SVM-based forecast structure by selecting the best parameters. The process is then executed in Weka and finally its output performance is determined using various matrices. Model determined as efficient is stored for later reference.

4. RESULTS AND ANALYSIS

We have developed the SVM based forecasting model in Weka tool by performing experiments over the dataset for every station. The forecast is done for the year 2016 in one year advance and it is represented as “*” from end of training data. Different metrics are used that perform relative measurements that compare the predicted result as the limit of the last known target value, that is WQ. The various matrices used are expressed as follows

1. Mean Absolute Error (MAE): sum(abs(forecasted - real)) / N
2. Mean Absolute Percentage Error (MAPE): sum(abs((forecasted - real) / actual)) / N

Where N denotes total number of instances.

3. Root Mean Squared Error (RMSE): sqrt(sum((forecasted - real)^2) / N)
4. Mean Squared Error (MSE): sum((forecasted - real)^2) / N

Forecasting Model Development for First Station “Devprayag” using SVM

For Devprayag, the proposed SVM model was developed. As seen in the table 2 the entries marked ‘*’ highlights the WQ forecast of Ganga River for the next twelve months from Jan. 2016 to Dec. 2016. The forecast-values of WQ at this station is revealed in the figure 4 where the water is shown to be in the category A all the way in 2016.

Table-2: Forecasted result: twelve months in advance for “Devprayag” using SVM

Time Stamp	Instance No.	WQ
July-2015	175	1
Aug-2015	176	1
Sept-2015	177	1
Oct-2015	178	1
Nov-2015	179	1

Dec-2015	180	1
Jan-2016	181*	0.9794
Feb-2016	182*	0.9657
Mar-2016	183*	0.9509
Apr-2016	184*	0.9362
May-2016	185*	0.9206
June-2016	186*	0.906
July-2016	187*	0.8909
Aug-2016	188*	0.8751
Sept-2016	189*	0.8606
Oct-2016	190*	0.8456
Nov-2016	191*	0.8295
Dec-2016	192*	0.8133

The performance of developed model is as in the table 3 which is computed using various measures that showed that the developed model attained a very poor accuracy of 19.07% (best among others) as the minimum MSE is 0.8093. Thus, we can conclude from the obtained results that the models developed using SVM for this station does not yields in a satisfactory forecast.

Table-3: The forecasting performance of SVM model for "Devprayag" Sampling Station

Performance Metric / **Target WQ**	**MAE**	**MAPE**	**RMSE**	**MSE**
1 step- ahead	0.6702	31.6384	0.8996	0.8093
2 step- ahead	0.7859	35.8692	1.0414	1.0845
3 step- ahead	0.8455	38.0018	1.1223	1.2595
4 step- ahead	0.8501	37.6022	1.1518	1.3266
5 step- ahead	0.861	37.7718	1.1754	1.3815
6 step- ahead	0.8485	37.0804	1.1733	1.3766
7 step- ahead	0.8385	36.5195	1.167	1.3618
8 step- ahead	0.8282	36.1739	1.1592	1.3437
9 step- ahead	0.8191	36.0057	1.1493	1.3208
10 step- ahead	0.8125	36.0081	1.1421	1.3045
11 step- ahead	0.8264	36.6102	1.1596	1.3446
12 step- ahead	0.8409	37.1145	1.1747	1.3799

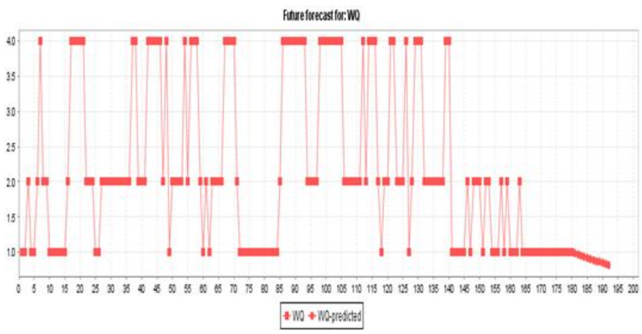

Fig. 4 Future Forecast for WQ (12-steps ahead) at Devprayag Station based on SVM

Forecasting Model Development for Second Station "Rishikesh" based on SVM

At the next station Rishikesh, the WQ for next twelve months is displayed in the table 4 whereas table 5 presents the forecasting performance of developed model. The forecast of WQ at this site is presented in the figure 5. Once again here the created model achieved an ideal accuracy of merely 26.48% with lowest MSE of 0.7352 only which is again not upto the mark as highlighted in table 5. Therefore, this is confirmed from the outcome that at this site also the model based on SVM failed again in forecasting the WQ.

Table-4: : Forecasted result: twelve months in advance for "Rishikesh" using SVM

Time Stamp	Instance No.	WQ
July-2015	175	1
Aug-2015	176	1
Sept-2015	177	1
Oct-2015	178	1
Nov-2015	179	1
Dec-2015	180	1
Jan-2016	181*	0.9973
Feb-2016	182*	0.9925
Mar-2016	183*	0.9855
Apr-2016	184*	0.9763
May-2016	185*	0.965
June-2016	186*	0.9516
July-2016	187*	0.9362
Aug-2016	188*	0.9188
Sept-2016	189*	0.8996
Oct-2016	190*	0.8787
Nov-2016	191*	0.8563
Dec-2016	192*	0.8324

Table-5: The forecasting performance of ANN model for station "Rishikesh" Sampling Station

Performance Metric / Target WQ	MAE	MAPE	RMSE	MSE
1 step- ahead	0.5139	24.7716	0.8574	0.7352
2 step- ahead	0.7265	35.4928	1.0741	1.1536
3 step- ahead	0.8947	43.3222	1.2176	1.4826
4 step- ahead	1.0053	48.9888	1.2922	1.6698
5 step- ahead	1.1098	55.2161	1.3705	1.8781
6 step- ahead	1.1709	60.0033	1.4095	1.9867
7 step- ahead	1.2111	63.2809	1.4278	2.0386
8 step- ahead	1.2243	64.8623	1.425	2.0305
9 step- ahead	1.2412	65.9377	1.4247	2.0298
10 step- ahead	1.2367	66.3982	1.4103	1.9889

11 step- ahead	1.2223	66.2455	1.3906	1.9339
12 step- ahead	1.2141	66.2535	1.3768	1.8957

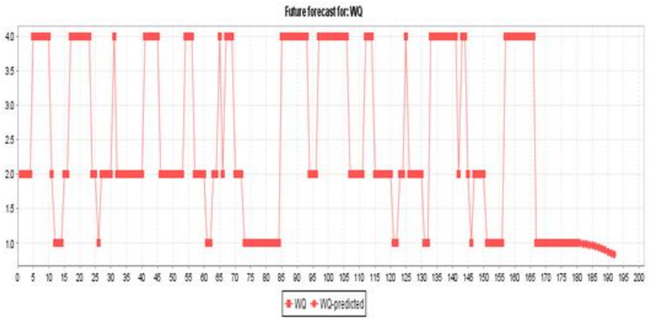

Figure 5: Future Forecast for WQ (12-steps ahead) at Rishikesh Station based on SVM

Forecasting Model Development for Third Station "Haridwar" based on SVM

The proposed model for Haridwar site, bring about in the forecast for the WQ of the Ganga River for the next twelve-months in advance as displayed in the table 6. and shown graphically in the figure 6. The model obtained a best accuracy of only 11.43% as emphasized in table 7 which is not acceptable performance. Therefore, once again the SVM technique failed to produce an efficient forecasting model for the undertaken problem.

Table-6: Forecasted result: twelve months in advance for "Haridwar" using SVM		
Time Stamp	**Instance No.**	**WQ**
July-2015	175	2
Aug-2015	176	2
Sept-2015	177	2
Oct-2015	178	1
Nov-2015	179	1
Dec-2015	180	1
Jan-2016	181*	0.944
Feb-2016	182*	0.8795
Mar-2016	183*	0.9023
Apr-2016	184*	0.9099
May-2016	185*	0.858
June-2016	186*	0.8258
July-2016	187*	0.795
Aug-2016	188*	0.7597
Sept-2016	189*	0.7055
Oct-2016	190*	0.669
Nov-2016	191*	0.631
Dec-2016	192*	0.5913

Table-7: The forecasting performance of SVM model for "Haridwar" Sampling Station				
Performance Metric	**MAE**	**MAPE**	**RMSE**	**MSE**

Target WQ				
1 step- ahead	0.6722	35.4457	0.9411	0.8857
2 step- ahead	0.8549	46.2695	1.1403	1.3003
3 step- ahead	1.0013	54.3924	1.2792	1.6364
4 step- ahead	1.0374	56.1609	1.3105	1.7175
5 step- ahead	1.0928	59.3242	1.3474	1.8156
6 step- ahead	1.096	59.5527	1.341	1.7982
7 step- ahead	1.0895	59.2013	1.3242	1.7535
8 step- ahead	1.106	60.0732	1.3352	1.7827
9 step- ahead	1.1078	60.9522	1.3492	1.8204
10 step- ahead	1.1245	62.7311	1.3758	1.8927
11 step- ahead	1.1411	64.6016	1.3952	1.9465
12 step- ahead	1.156	66.3453	1.4174	2.0091

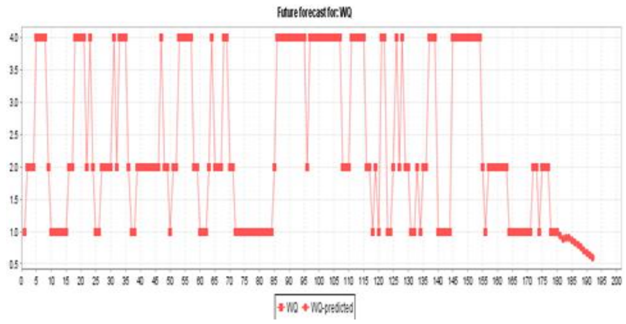

Fig. 6 Future Forecast for WQ (12-steps ahead) at Haridwar Station

Forecasting Model Development for Fourth Station "Jwalapur"

At station Jwalapur, the water quality of the Ganga River is displayed in the table 8 and performance of model is declared in the table 9. The forecast of water quality at this station is presented as a graph in the figure 7. But as usual we don't belief and accept these results because the developed model achieved an ideal forecasting accuracy of just 13.47%.

Table-8: Forecasted result: twelve months in advance for "Jwalapur" using SVM

Time Stamp	**Instance No.**	**WQ**
July-2015	175	4
Aug-2015	176	4
Sept-2015	177	4
Oct-2015	178	4
Nov-2015	179	4
Dec-2015	180	4
Jan-2016	181*	4.1201
Feb-2016	182*	4.168
Mar-2016	183*	4.2026
Apr-2016	184*	4.2239
May-2016	185*	4.2412
June-2016	186*	4.2481

July-2016	187*	4.2558
Aug-2016	188*	4.2562
Sept-2016	189*	4.2589
Oct-2016	190*	4.2585
Nov-2016	191*	4.2589
Dec-2016	192*	4.2575

Table-9: The forecasting performance of SVM model for "Jwalapur" Sampling Station

Performance Metric / Target WQ	**MAE**	**MAPE**	**RMSE**	**MSE**
1 step- ahead	0.6604	34.3961	0.9302	0.8653
2 step- ahead	0.7514	39.7152	1.0287	1.0581
3 step- ahead	0.7861	41.6801	1.0731	1.1516
4 step- ahead	0.7919	42.3458	1.0882	1.1841
5 step- ahead	0.8104	43.2944	1.1069	1.2253
6 step- ahead	0.8153	43.5105	1.1093	1.2305
7 step- ahead	0.8174	43.7135	1.1087	1.2291
8 step- ahead	0.814	43.727	1.1051	1.2213
9 step- ahead	0.8163	44.1672	1.1099	1.2318
10step- ahead	0.8235	44.8653	1.1193	1.2529
11step- ahead	0.8389	45.8502	1.1368	1.2922
12step- ahead	0.8513	46.7658	1.1549	1.3337

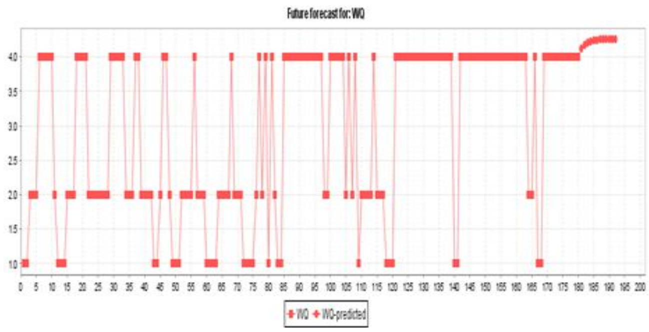

Fig. 7 Future Forecast for WQ (12-steps ahead) at Jwalapur Station

Forecasting Model Development for Fifth Station "Roorkee"

Finally at the fifth sampling station Roorkee the developed model indicated forecast for quality of water of the Ganga River in advance for upcoming twelve months in the table 10 and the forecast-values in the figure 8. It is determined from these results that though the quality of water at this site forecasted to be part of category A but the model accomplished a superlative accuracy of only 39.39% in forecasting one month advance as highlighted in table 11 which is not a satisfactory forecasting performance.

Table-10: Forecasted result: twelve months in advance for "Roorkee" using SVM

Time Stamp	Instance No.	WQ
July-2015	175	1
Aug-2015	176	1
Sept-2015	177	1
Oct-2015	178	1
Nov-2015	179	1
Dec-2015	180	1
Jan-2016	181*	0.9709
Feb-2016	182*	0.958
Mar-2016	183*	0.9465
Apr-2016	184*	0.9352
May-2016	185*	0.9248
June-2016	186*	0.9164
July-2016	187*	0.9091
Aug-2016	188*	0.9017
Sept-2016	189*	0.8944
Oct-2016	190*	0.8851
Nov-2016	191*	0.8771
Dec-2016	192*	0.8694

Table-11: The forecasting performance of SVM model for "Roorkee"

Performance Metric / Target WQ	MAE	MAPE	RMSE	MSE
1 step- ahead	0.5198	25.1821	0.7785	0.6061
2 step- ahead	0.5991	28.5736	0.8897	0.7916
3 step- ahead	0.6408	29.8333	0.9501	0.9028
4 step- ahead	0.6553	29.8725	0.9935	0.987
5 step- ahead	0.6608	30.2307	1.0159	1.032
6 step- ahead	0.6581	30.1464	1.0196	1.0396
7 step- ahead	0.6512	29.9179	1.016	1.0322
8 step- ahead	0.6439	29.6694	1.0129	1.0259
9 step- ahead	0.6349	29.3997	1.0072	1.0144
10 step- ahead	0.6318	29.2393	1.0129	1.0259
11 step- ahead	0.6406	29.5487	1.0214	1.0432
12 step- ahead	0.6533	30.0646	1.0326	1.0663

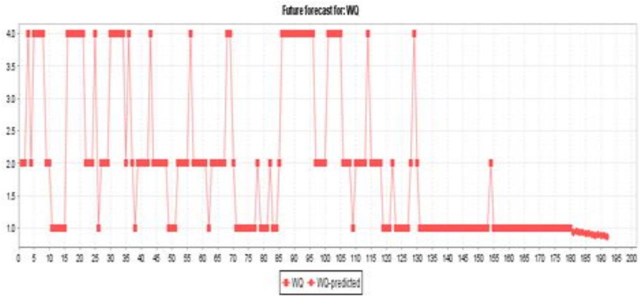

Fig. 8 Future Forecast for WQ (12-steps ahead) at Roorkee Station

Therefore, even SVM-based technology for this station failed to develop a forecasting model that could efficiently predict the water quality of the Ganges River.

5. CONCLUSION

In an effort to develop a WQ forecasting model for the Ganges River using the technique of SVM, the results lead to the conclusion that the SVM-based WQ forecasting model has not performed efficiently for the current dataset. However, if we compare these results to the water quality forecasts obtained through ANN, as published in [1], we found that ANN outperforms SVM in forecasts.

It is noted that with this paper the SVM achieves the best model accuracy of only 39.39% while using ANN we found 100% accuracy [1]. Therefore in the future, further attention will be paid to improving this current water quality forecasting model for a river based on SVM, by including another configuration for the model such as changes in kernel function.

6. REFERENCES

[1] Anil Kumar Bisht, Ravendra Singh, Ashutosh Bhatt, Rakesh Bhutiani, "Artificial Neural Network Based Water Quality Forecasting Model for Ganga River", International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Volume-8, Issue-6, August 2019, PP: 2778-2785.

[2] Yakup Kara, Melek Acar Boyacioglu, Omer Kaan Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange", ELSEVIER, Expert Systems with Applications 38 (2011) pp. 5311–5319.

[3] Amruta A. Taksande , P. S. Mohod, "Applications of Data Mining in Weather Forecasting Using Frequent Pattern Growth Algorithm", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064, Volume 4 Issue 6, June 2015.

[4] Srinivas Mukkamala, Andrew H. Sung, Ajith Abraham, "Intrusion detection using an ensemble of intelligent paradigms", Journal of Network and Computer Applications, ELSEVIER, 28 (2005) 167–182.

[5] Nattane Luiza Costa, Laura Andrea Garcia Llobodanin , Inar Alves Castro, Rommel Barbosa, "Using Support Vector Machines and neural networks to classify Merlot wines from South America", Journal of Information Processing In Agriculture, ELSEVIER, 6 (2019) 265–278.

[6] Hamed Ahmadi, Markus Rodehutscord, "application of artificial neural network and support Vector Machines in Predicting Metabolizable energy in compound Feeds for Pigs", Journal of Frontiers in Nutrition,Volume 4, June 2017.

[7] Ehsan Olyaie, Hamid Zare Abyaneh, Ali Danandeh Mehr, "A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River", ELSEVIER, Geoscience Frontiers xxx (2016) 1-11.

[8] Huihui Yu, Yingyi Chen, ShahbazGul Hassan, Daoliang Li, "Dissolved oxygen content prediction in crab culture using a hybrid intelligent method", Scientific RepoRts | 6:27292 | DOI: 10.1038/srep27292, www.nature.com/scientificreports, 2016.

[9] Liao, Y., Xu, J., & Wang, W. (2011). A Method of Water Quality Assessment Based on Biomonitoring and Multiclass Support Vector Machine. Procedia Environmental Sciences, Elsevier, 10, 451–457. doi:10.1016/j.proenv.2011.09.074
[10] http://www.cs.waikato.ac.nz/ml/weka.
[11] www.mathworks.com
[12] Yue Liao, Jian-yu Xu, Zhu-wei Wang, "Application of biomonitoring and support vector machine in water quality assessment", Journal of Zhejiang University (Biomed & Biotechnol), Vol. 13(4):pp. 327-334. ISSN 1673-1581 (Print); ISSN 1862-1783 (Online), 2012.

AUTHORS PROFILE

Mr. Anil Kumar Bisht is currently a Ph.D. research scholar and Asstt. Professor in the department of CS & IT, FET, Mahatma Jyotibha Phule Rohilkhand University, Bareilly, U.P. India. He has a teaching experience of 16 years. He has done his M.Tech. in CSE from Uttarakhand Technical University, Dehradun, Uttarakhand, India. He has done his B.Tech in CSE from Uttar Pradesh Technical University, Lucknow, U.P., India. His research Areas are Artificial Intelligence, Machine Learning and Wireless Communications. He has completed online certificate course from Electronics and ICT Acedemy, IIT Kanpur on "Machine Learning". He has a total number of 18 reputed research publications. He is a life time member of Computer Society of India (CSI).

Dr. Ravendra Singh is a Professor in the department of CS & IT, FET, Mahatma Jyotibha Phule Rohilkhand University, Bareilly, U.P. India. He has a teaching experience of 22 years and 5 years experience in industry as Assistant Manager in Uptron India Ltd., Lucknow (U.P. Govt. Undertaking). His research Areas are at present, mainly engaged in Artificial Intelligence, Wired & Wireless networks, QoS based routing, ad hoc network, VANET, routing in under water and terrestrial sensor network, Admission control schemes for Real time communication and guided many scholars over the Internet, simulation and performance evaluation of wireless networks, Cost optimization & testing issues and Task allocation & scheduling in Distributed parallel computers. He is also working in the area related to association rules, cleaning & Integration of data in Data mining, & Data warehousing. He has a total number of more than 37 reputed publications. Also he had written 2 books with New Age International, N.Delhi and University Science Press, N.Delhi. He is a member of AICTE, NAAC and other government academic bodies in India.

Dr. Rakesh Bhutiani is currently working as Assistant Professor in deptt. of environmental science, Gurukul Kangri University, Haridwar (UTTARAKHAND) INDIA. He has done his Ph.D. in Environmental Science. He has a total 18 years of teaching and research experience. He has more than 80 reputed publications in National and International Journals of repute and guided several scholars. Dr. Bhutiani is fellow of several academic societies and has several awards to his credit.

Dr. Ashutosh Bhatt is in Deptt. of CSE at Birla Institute of Applied Sciences, Bhimtal, Uttarakhand, India. He has done his Ph.D. in CSE in the field of Machine Learning. He is having a teaching experience of 22 years. He has a good number of reputed research publications. His Interest area are: Neural Networks, Web Applications. He has guided many research scholars. He is a member of different research societies.