IOT BASED PREDICTION OF WATER QUALITY INDEX FOR FARM IRRIGATION

CREATED BY:

Dr. RAJESH KUMAR YADAV

ADARSH JHA

ADITYA CHOUDHARY

INTRODUCTION

- Less Productivity in Agriculture Sectors(50 % workforce contributes 16% to the GDP). There exists a need to increase the efficiency of each stage in farming and at a cheap cost so that it is affordable. With improvement in technology, these needs can be addressed using innovative solutions like the Internet of Things(IoT).

- Salinity of water used in irrigation can affect soil nutrients and plant growth as described in the works of Mark[1] and Shrivastav[2]. Based on this, Meireles developed an Irrigation Water Quality Index(IWQI)[3].

- In this paper we focus on determining irrigation water quality by developing a classification model for prediction of IWQI class based on IWQI developed by Meireles.

[1] M. E. Grismer and K. M. Bali, “Drought tip: Use of saline drain water for crop production,” 2015
[2] P. Shrivastava and R. Kumar, “Soil salinity: a serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation,” Saudi journal of biological sciences, vol. 22, no. 2, pp. 123–131, 2015.
[3] A. C. M. Meireles, E. M. d. Andrade, L. C. G. Chaves, H. Frischkorn, and L. A. Crisostomo, “A new proposal of the classification of irrigation water,” Revista Ciencia Agronomica, vol. 41, no. 3, pp. 349–357, 2010

OBJECTIVES

- ❖ We convert 5 parameters: Na^+, Cl^-, EC, HCO^{3-} and SAR to quality measurement values which are used to calculate IWQI. We construct IWQI classes using IWQI values for irrigation water.

- ❖ We reduce 5 parameters to 3 parameters using correlation analysis to save costs of sensors required to implement the proposed work. These parameters are used to classify IWQI using seven classification techniques which are Support Vector Classifier, Neural Networks, Gradient Boosting , Random Forest, Decision Tree, Bagging and Naive Bayes classifier.

- ❖ We evaluate and choose the best performing classification algorithm to obtain IWQI class for water sample with 3 parameters.

IRRIGATION WATER QUALITY INDEX

- The proposed IWQI is an aggregation of 5 parameters: Na^+, Cl^-, EC, HCO^{3-} and SAR.

- These 5 parameters are measured and converted to quality measurement values (q_i values) because
  - All these parameters have different units
  - Water quality decreases if values of the parameters are outside the threshold limits.

- After obtaining quality measurement values, IWQI is calculated using the formula: $IWQI = \sum q_i * w_i$[4] where w_i is the relative weight of the parameters. The values of IWQI are in range 0 – 100 which can be divided into classes as shown in Table 1

[4] Abbasnia, Abbas, et al. "Evaluation of groundwater quality using water quality index and its suitability for assessing water for drinking and irrigation purposes: Case study of Sistan and Baluchistan province (Iran)." *Human and Ecological Risk Assessment: An International Journal* 25.4 (2019): 988-1005.

Table 1: Range for IWQI classes

IWQI	Soil	Plant
85–100	Can be used for any kind of soil	Most plants won't be affected
70–85	Can be used on soil with moderate permeability	Avoid use in plants with very low salt tolerance
55–70	Can be used on soils with moderate to high permeability	Avoid in plants with low salt tolerance
40–55	Can be used on soils with high permeability without dense layers	Used mainly in plants with high salt tolerance. Plants with moderate salt tolerance can be used with some control practices
0–40	Use for irrigation should be avoided	Avoid for all plants

Dataset: Major Ions Dataset

The dataset used in our work was obtained from US Geological Survey of Brackish Groundwater [5]. We select the parameters required to calculate IWQI which are mentioned in Table 2 along with their units.

Sodium, Magnesium and Calcium ions concentrations are used to calculate SAR.

Table 2: IWQI parameters

Name of parameter	**Unit**
Sodium ion	mmol/L
Calcium ion	mmol/L
Magnesium ion	mmol/L
Chloride ion	mmol/L
Electrical conductivity	dS/cm
Bicarbonate ion	mmol/L

[5] S. Qi and A. Harris, "Geochemical database for the brackish groundwater assessment of the united states," US Geological Survey data release, 2017.

DEVELOPING THE CLASSIFICATION MODEL

We calculate IWQI using 5 parameters from Major ions dataset and convert them to IWQI classes according to ranges given in Table 1.

We perform correlation analysis on the parameters in IWQI which is given in Table 3. The following observations are made:

1) Cl^- is highly correlated with EC and Na^+.
2) Na^+ is highly correlated with all the parameters.
3) EC is highly correlated with all the parameters except SAR.
4) HCO^{3-} is highly correlated with EC.
5) IWQI is highly correlated with Cl^-, Na^+ and EC.

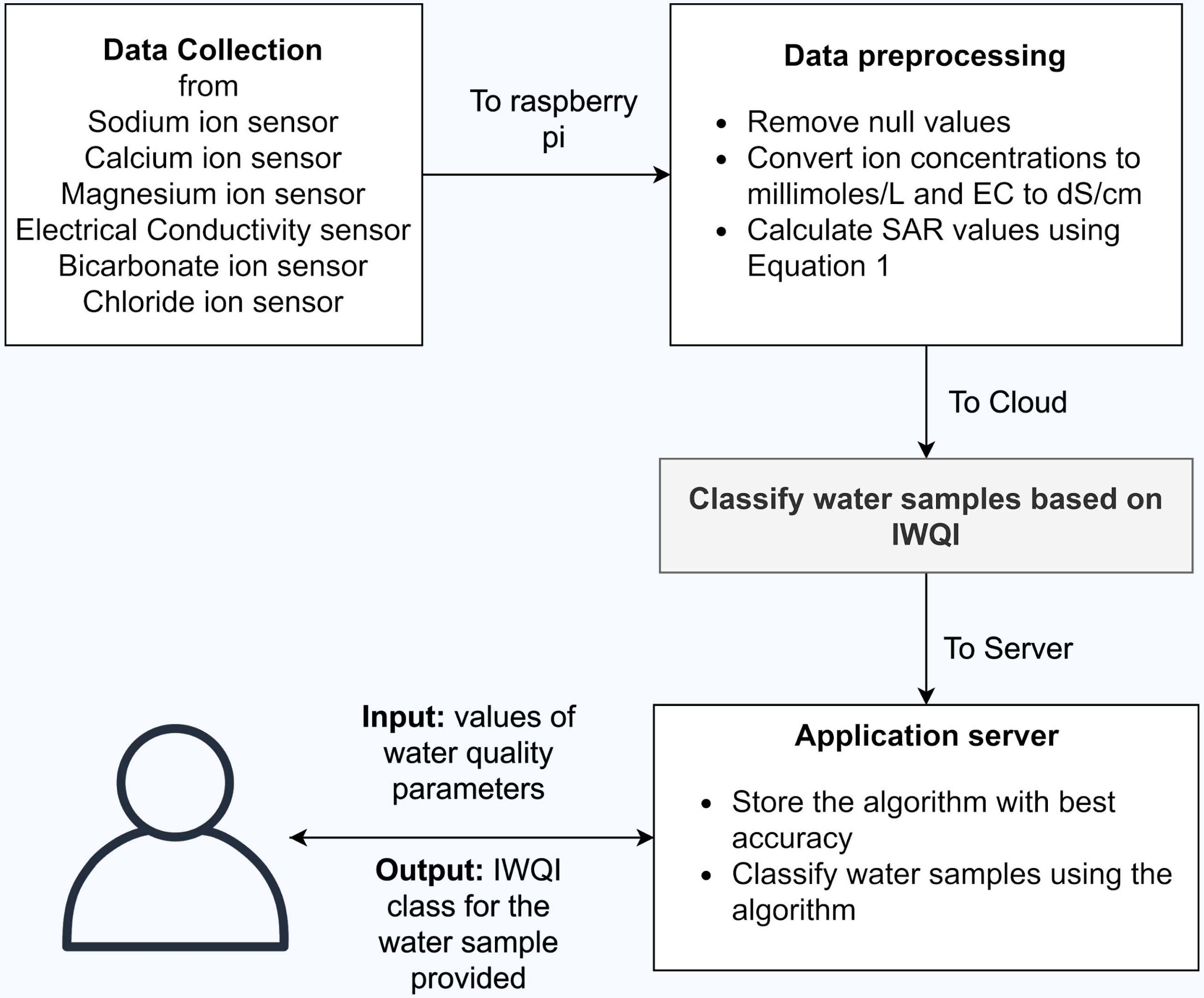
We select 3 parameters: Cl^-, Na^+ and EC according to their correlation with IWQI as features for the classification algorithm.

Table 3: Correlation matrix for various parameters for water samples

	Cl^-	**Na^+**	**EC**	**HCO^{3-}**	**SAR**	**IWQI**
Cl^-	1	0.467	0.506	0.159	0.083	0.44
Na^+	0.467	1	0.482	0.297	0.56	0.607
EC	0.506	0.482	1	0.835	0.018	0.423
HCO^{3-}	0.159	0.297	0.835	1	0.018	0.293
SAR	0.083	0.56	0.018	0.018	1	0.302
IWQI	0.44	0.607	0.423	0.293	0.302	1

To cover variety of classification methods, 7 classification algorithms are used which are: Support Vector Classifier, Neural Networks, Gradient Boosting , Random Forest, Decision Tree, Bagging and Naive Bayes classifier.

OVERALL ARCHITECTURE



EVALUATION METRICS

The classification algorithms are evaluated using the following metrics:

1) Accuracy[6]: It tell us how much data can be predicted accurately from a given classifier. This gives us an idea about the bulk of data which is correctly classified.
2) Precision[6]: It tells us how many predictions are correct out of all observations saying output belongs to a particular class.
3) Recall[6]: It tells us how many correct predictions were made for a class out of all the actual observations for a given class.
4) F1 Score[6]: It can be defined as harmonic mean of precision and recall, can be used independently for evaluation of classifiers.

[6] Powers, David MW. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." arXiv preprint arXiv:2010.16061 (2020).

RESULTS

Random Forest performed the best with an accuracy of 86.9% followed by Gradient Boosting and Neural Networks with accuracy of 85.8% and 84.6% respectively. Naive Bayes classifier performed the worst with accuracy of 52.6%. The classification results and cross validation scores are given in Table 5 and 6 respectively.

Table 5: Classification results for 7 classifiers

Methods	**Accuracy**	**Precision**	**Recall**	**F1**
DecisionTree	0.832	0.728	0.740	0.733
Naive Bayes	0.526	0.105	0.200	0.138
Gradient Boosting	0.858	0.762	0.757	0.760
Random Forest	0.869	0.790	0.765	0.776
SVM	0.845	0.762	0.711	0.726
Bagging	0.813	0.750	0.664	0.690
MLP	0.846	0.734	0.743	0.738

Table 6: Cross validation scores for 7 classifiers

Methods	**Fold1**	**Fold2**	**Fold3**	**Fold4**	**Fold5**
Bagging Classifier	0.829	0.823	0.814	0.830	0.813
DecisionTree	0.833	0.836	0.828	0.823	0.839
Naive Bayes	0.518	0.518	0.518	0.518	0.519
Gradient Boosting	0.862	0.866	0.869	0.859	0.862
Random Forest	0.866	0.870	0.870	0.862	0.871
SVM	0.853	0.849	0.846	0.841	0.848
MLP	0.847	0.853	0.857	0.847	0.849

CONCLUSION AND FUTURE WORKS

- In this work, we developed a classification model for prediction of Irrigation Water Quality Index (IWQI) class.

- For classification, we used seven classification algorithms and selected three out of five parameters: EC, Cl^- and Na^+ for predicting IWQI class. Random Forest Classifier performed the best followed by Gradient Boosting and Neural Networks.

- Other parameters such as acidity, oxygen demand etc. can be used in addition to the parameters used here to develop an index covering more factors. Also, dataset covering different types of water bodies can be used to improve the model.

- The proposed work used here can be incorporated into IoT based agricultural systems for analysing irrigation water quality that will be faster and economically feasible compared to manual lab tests. It can also help in deciding the water sample according to crop and soil properties.

REFERENCES

[1] R. S. Ayers, D. W. Westcot and others, Water quality for agriculture, vol. 29, Food and Agriculture Organization of the United Nations Rome, 1985.

[2] L. Breiman, "Bagging predictors," Machine learning, vol. 24, p. 123–140, 1996.

[3] L. Breiman, "Random forests," Machine learning, vol. 45, p. 5–32, 2001.

[4] V. Chinnusamy, A. Jagendorf and J.-K. Zhu, "Understanding and improving salt tolerance in plants," Crop

science, vol. 45, p. 437–448, 2005.

[5] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Annals of statistics, p.

1189–1232, 2001.

[6] R. Gentili, R. Ambrosini, C. Montagnani, S. Caronni and S. Citterio, "Effect of soil pH on the growth, reproductive investment and pollen allergenicity of Ambrosia artemisiifolia L.," Frontiers in plant science, vol. 9, p.

1335, 2018.

[7] R. Gentili, R. Ambrosini, C. Montagnani, S. Caronni and S. Citterio, "Effect of Soil pH on the Growth, Reproductive Investment and Pollen Allergenicity of Ambrosia artemisiifolia L.," Frontiers in Plant Science, vol. 9, p. 1335, 2018.

[8] M. E. Grismer and K. M. Bali, "Drought Tip: Use of Saline Drain Water for Crop Production," 2015.

[9] B. Gupta and B. Huang, "Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization," International journal of genomics, vol. 2014, 2014.

[10] M. H. Hassoun and others, Fundamentals of artificial neural networks, MIT press, 1995.

THANK YOU