Nandha Kumar G

| | | |
|---|---|---|
| Each data needs to be in different measures to analyze the quality | Prediction can also be taken from the historical dataset | Keep the data design |
| The proposed prediction system will iteratively test the model with training and testing datasets | Data modeling to use the past dataset to inform the future effort | The data mining techniques will be used for applying the classification method for water quality application |

Nalinashree N

| | | |
|---|---|---|
| The data distribution in the testing data should not affect the training data set. | Various techniques can be included to predict the quality within the application. | Use a minimal number of parameters with cheap sensors to predict water quality |
| Using supervised learning algorithm, water quality class can be predicted | Cross-validation can used to evaluate method for reducing scales of overfitting and increasing accuracy of the model | Variable importance analysis can increase the accuracies of the models |

Renuga devi N

| | | |
|---|---|---|
| Massive dataset and strong correlation between parameters will make the best prediction. | Accurate model can be selected based on the outcome in the model evaluation | Network structure selection method is proposed to identify the corelated input parameters |
| A method like neuro-fuzzy interference system can be implemented which is capable of integrating linear and non-linear relationships in dataset. | Evaluating the effect of substantial nutrient loads on overall water quality | Some of the variables can be eliminated due to the meaningless analysis |

Vikram S

| | | |
|---|---|---|
| The size of training datasets should not be less than the number of training parameters required in the model. | Stratified sampling strategy is used to mitigate the uneven distribution of training and testing dataset | The timeline of the measurements must be recorded |
| Parameters like temperature, turbidity, pH and dissolved solids can be used | Feature selection helps to simplify the procedure and reduce computational cost of analysis | The variable importance measure must be weighted sums of the absolute regression coefficients. |