A WEB BASED VISUALIZATION OF WATER QUALITY CORRELATION WITH ENVIRONMENTAL INFLUENCE USING CLOUD PLATFORM

Abstract

Water quality variation in river is occurring the result of human behavior, and natural processes. The most possible solution to restraint the water contamination in river is recognize and control the anthropogenic activities. To overcome this problem, improve self-discipline by visualizing effect of human behavior is proposed as a sophisticated solution. This research work discusses about comparative study for development of the model and informative web interface to visualize water quality correlation with environmental factors based on cloud computing.

Introduction

Water quality classification of a particular river basin describes the water quality status of known locations. It is a scientific method to identify suitable locations to fulfill the requirements of people such as, drinking, bathing, agriculture purposes and other activities. How can we use water quality classification model for evaluate the effect of anthropogenic activities? Therefore in here evaluated the effect of population for Total coliform level in a river basin. The following mentioned materials in Table 1 and Table 2 were used to obtained models and selected sub streams are shown in Fig 1.

Table 1: Selected classification levels from Proposed water quality classification (Western Region Megapolis Planning (unpublished, 2015))

<table>
<thead>
<tr>
<th>Class</th>
<th>Level X &lt;= 5000</th>
<th>X &lt;= 20,000</th>
<th>X &gt; 20,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Evaluated Issues

• Can we drink water. When Total coliform is XXX & Population is XXXX?

\[ P(B|TC = 800, Population = 18,850) = \frac{P(B|TC = 800) \cdot P(B)}{P(TC = 800)} \]

• How can we communicate each other when we need to establish water quality management system for River basin?

Methodology

Data set.

138 Total coliform data and corresponding population in 2001 and 2012.

Development of the ANN Model

Then optimization of performance of model was conducted in two main stages. In the first stage the impact of number of hidden neurons and then optimized the training time on network’s test correctness in the second stage.

Development of the BN Model

Accuracy of model was evaluated for 1, 2, 3 different maximum number of parents of K2 global score metric based learning algorithm with default setting of simple estimator.

Model evolution technique.

The cross validation technique. The accuracy was calculated by 1/3 of test dataset after training the model with 2/3 of the training sample. The values of percentage of correctly classified, Kappa statistic J-measure and computational time are considered to evaluate the performance of model in each stage.

Result and Discussion

Given quality classification in Table 1, drinking water category overlaps with bathing water category and both these overlap with fishing category.

Classification Model with ANN

1st Step—The result for test correctness evaluation by increasing training time.

2nd Step—The result of increasing number of hidden neurons by constant number of epochs.

3rd Step—The result of increasing number of fold of cross validation by constant number of epochs 1,000 and 10 hidden neurons.

The maximum percentage of correctly classified data records reached 89.02% for 10 hidden neurons, 1,000 epochs in 10 fold cross validation.

Classification Model with BN

Increasing the number of parent of K2 learning algorithm, evaluated the performance of the BN model. The BN model obtained best performance 100% of correctly classified data records reached.

Considering only population attribute with classification class

Some procedure was followed as mention in above

Result

ANN—Accuracy rate = 51.44%; Kappa statistic = 0.121

BN—Accuracy rate = 46.24%; Kappa statistic = 0.0

Data set

• Some population is considered by all records (12-months) in particular period.
• Data is consist with 138 records only.

BN

• BN is good classification than ANN according to this situation. In 1st comparison, BN directly followed the TC not consider population, Even it has high accuracy than ANN.