

# Red Tide Forecast using Ensemble Method

Sun Park, Soonyeong Park, Ilwhan Oh, Gyeongho Kim, Jonghwan Park,  
Seong Ro Lee

Mokpo National University, South Korea  
{sunpark, sypark, ilwhan, khkim, jongpark, srlee}@mokpo.ac.kr

**Abstract.** To enhance the automatic forecast of red tide, this paper proposes a red tide prediction method that uses ensemble method. The proposed method improves the precision of forecast results because the ensemble classifier is enhanced by optimal data of the proposed preprocessing.

**Keywords:** Red Tide, forecast, ensemble.

## 1 Introduction

In recent year, red tide has been studied extensively according to the increment of a sea farming damage by means of red tide blooms in connection with harmful algae in the world. Research on red tide events of South Korea has mostly focused chemical properties and investigation of biological causes, which is generally used for mitigation activity. However, there has been minimal research regarding variety of automatic methods with relation to monitoring and predicting red tide blooms in the coast of South Korea [1, 2]. Though automatic forecast of red tide events has many limitations, a red tide bloom can be automatically predicted by analyzing past data of the occurrence of red tide and current data of the marine environment. Recently, other techniques for red tide forecast, including the manual method [3], case-based reasoning (CBR) [4], and pattern recognition [5] have been proposed. This paper proposes a red tide forecast method using ensemble method.

## 2 Proposed Method

The proposed method consists of two phases: preprocessing, ensemble method. In the following subsections, we explain each phase in full.

### 2.1 Preprocessing

In this section, the input and output variables for preprocessing phase with regard to the classifier is modeled by using the past red tide event data and the marine environmental data. The input variables used are the water temperature, surface sea

temperature, and rainfall. The output variables use the density of red tide algae and the red tide bloom. The range of input variables is initialized 10 days before the prediction of the red tide event.

The input and output variables of the preprocessing phase are described as follows. The input variables of water temperature and surface sea temperature are defined by the average of 10 days before the red tide blooms. The input variable of rainfall is defined by the sum of 10 days before the red tide event.

## 2.2 Ensemble method

In the ensemble method phase, the categorical red tide forecast is predicted by bagging based on decision tree [6]. The ensemble algorithm uses the bagging algorithm since it also shows better prediction performance than the boost algorithm in our training and test data. The ensemble method phase consists of the choice of basis classifier and the bagging method.

## 3 Experiments and Evaluation

In this paper, we used the experimental data of 1,461 actual events of harmful algal of *cochlostinium p.* in the nine years from 2002 to 2010 that occurred on the coast of South Korea. The data was gathered from the red tide news bulletin of the red tide information system [1]. In addition, marine environmental data in connection with the data of the red tide events was collected (i.e., the temperature and rainfall from Korea National Weather Service [7] and the water temperature from the Korea National Fisheries Research & Development Institute [8]). The training data used to construct the classifier using the six years from 2002 to 2007, and the test data used to evaluate the prediction result by the four years from 2007 to 2010.

In this paper, we use five different prediction methods: BPNN, SVM, C45, CART, and ESB. BPNN denotes the back propagation neural network [6]. SVM denotes the support vector machine [6]. C45 [6] and CART [6] denote decision tree algorithms. ESB is the proposed method. The performance evaluation was conducted by comparing the results of proposed method with the results of four prediction classifiers using the same data including the preprocessed and original data. The average accuracy of ESB is 19.57% higher than that of BPNN, 4.27% higher than that of SVM, 3.59% higher than that of C45, and 2.23% higher than that of CART. ESB shows best performance.

## 4 Conclusion

This paper proposed a red tide forecast method using ensemble method to enhance the prediction results. The preprocessing is proposed to model variable data to improve the classifier performance. The ensemble model is proposed to enhance the categorical prediction result.

## Acknowledgement

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0093828). "This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)"(NIPA-2012-H0301-12-2005).

## References

1. Red Tide Information System, Korea National Fisheries Research & Development Institute, "<http://portal.nfrdi.re.kr/redtide/index.jsp>", (2012)
2. Y. M. Kim, Y. G. Byun, Y. Huh, K. Y. Yu, "Detection of *Cochlodinium Polykrikoides* Red Tide Using MODIS Level 2 Data in Coastal Waters", *Journal of Korean Society of Civil Engineers*, pp. 535-540, (2007)
3. J. K. Kim, H. J. Youn, "HABs Prediction method in Yeosu bay using Remote Sensing", In proceeding of spring Conference of Korean Society GIS, pp.47-57, (2001)
4. B. H. Song, M. A. Jung, S. R. Lee, "A Design and Implementation Red Tide Prediction Monitoring System using Case Based Reasoning", *Journal of Korea Information and Communications Society*, vol. 35 no.12, pp.1819-1826, (2010)
5. Z. Rong, Y. Hong, D. Liping, "Research on Prediction of Red Tide Based on Fuzzy Neural Network", *Marine Science Bulletin*, vol. 8 no. 1, (2006)
6. J. Han, M. Kamber, "Second Edition, *Data Mining Concepts and Techniques*", Morgan Kaufmann Publishers, (2006)
7. Korea National Weather Service, "<http://www.kma.go.kr/index.jsp>", (2013)
8. Korea National Fisheries Research & Development Institute, "<http://portal.nfrdi.re.kr/>", (2013)