AN IMPROVED ENSEMBLE APPROACH FOR REDUCTION OF FALSE ALARM RATE IN HARMFUL ALGAL BLOOM DETECTION

Balakrishna Gokaraju, Surya S. Durbha, Roger L. King, and Nicolas H. Younan
Department of Electrical and Computer Engineering
GeoSystems Research Institute (GRI)
Mississippi State University, Mississippi State, MS 39762-9571, USA
balu@gri.msstate.edu, suryad@cavs.msstate.edu, rking@cavs.msstate.edu, and younan@ece.msstate.edu

Abstract--Harmful Algal Blooms (HABs) pose an enormous threat to U.S. marine habitation and economy in coastal waters. Federal and state coastal administrators have been working in devising a state-of-the-art monitoring and forecasting system for these HAB events. These modernized HAB systems provide useful and forewarning information to a varied user community. The current approaches are based on optical detection techniques which tend to give high false alarm rate in HAB detection. The chlorophyll anomaly method tends to confuse chlorophyll rich *Trichodesmium spp* and spring diatom Non-HABs with *K.brevis* HABs, and tends to give more false positives. Though, the band-ratio detection methods show a slight improvement in reducing the false positives, they tend to miss the target HABs (false negatives).

In the evaluation experiments of these approaches, the detection accuracy of HABs is found to be unsatisfactory due to the huge false alarm rate; hence a need for combined heuristic models or ensemble methods is envisaged [1]. These optical detection approaches are based on empirical relationships and use certain threshold values to classify HABs and Non-HABs. However, to improve upon false alarm rate, the detection approaches should employ learning techniques in their classification schemes and aim for better tuning to obtain a more complex and generalized models. We also need an approach employing the combination effect of these detection algorithms because, the optical detections can address the issues of classifying HABs from Non-HABs whereas the MODIS-FLH product with certain threshold value can address the classification issues of CDOM dominant waters from HABs. Thus, we need a linear combination of some model fitting technique instead of using a single fit of the technique. These techniques improve the predictive performance of a given statistical learning or a model fitting method. Ensemble method is one of the techniques that combines these detection algorithms and learning algorithms, and predicts HABs with high confidence by reducing the false alarm rate.

Ensemble methods combine multiple trained component classifiers' predictions when classifying new instances. These are generally believed to be more accurate than any of the component classifier. The two popular ensemble methods are Bagging and Boosting. These methods rely on resampling techniques to obtain different training sets for each of the classifiers. In this paper, we present an improved Ada-boosting ensemble approach of different component classifiers with neural network decision fusion for HAB detection. In this work, the ensemble of different classifiers is used to devise a better model to address various false alarm issues and generalization capability. The machine learning algorithm employed here is support vector machines that achieves good generalization. The ensemble methodology adopted in this approach is shown in Figure 1.

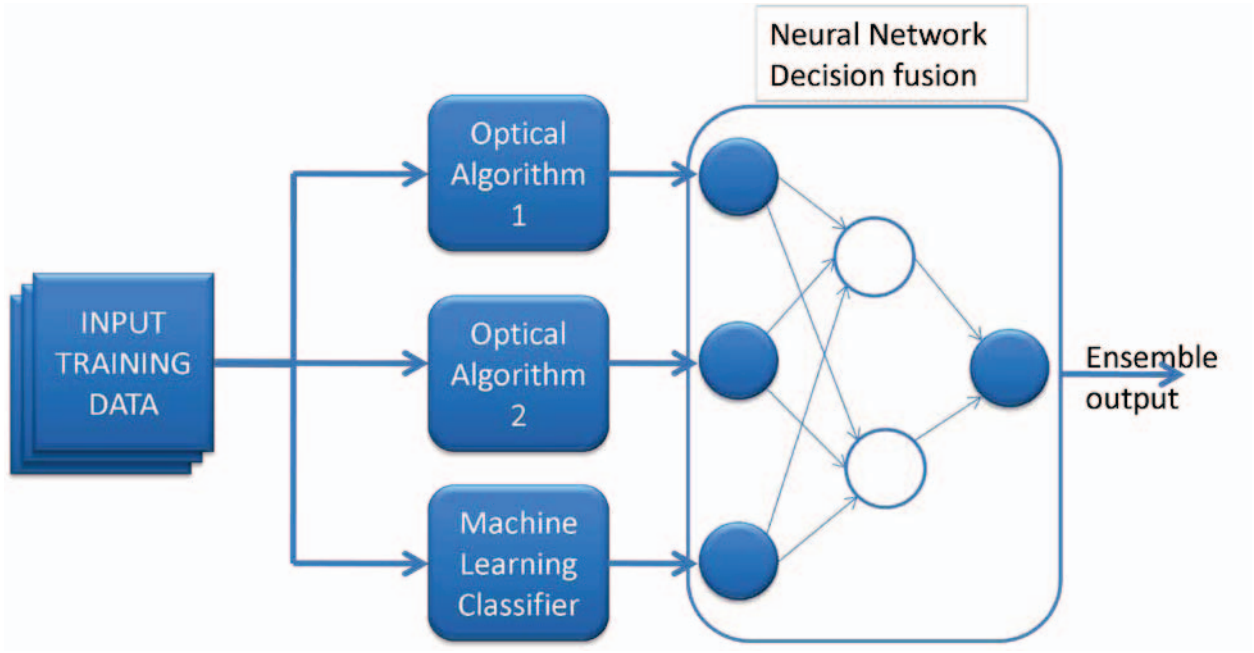

Figure 1: Ensemble Classifier with Neural Network Decision Fusion.

The efficiency of ensemble method relies on combining the individual component classifier's predictions. This approach fuses a neural network in place of weighted voting on the classifier's predictions. The neural network weighted voting gives more accurate predictions than the conventional probabilistic weighted voting. Here the weight distribution can be finely tuned with

the help of back propagation neural network learning. This modified ensemble approach helps in improving the HAB detection performance by reducing the false alarm rate and by obtaining generalization. The performance measure analyses are carried out using the N-fold cross validation accuracy, precision, recall and F-measures. We exhaustively compare various classifier combinations in the ensemble to find the best classifier combination for enhanced HAB detection.

REFERENCES

[1]. M. C. Tomlinson, T. T. Wynne and R. P. Stumpf, "An Evaluation of remote sensing techniques for enhanced detection of the toxic dinoglagellate, *Karenia brevis,*" *Remote Sensing of Environment*, Vol. 113, No. 3, pp. 598-609, 2009