CSI5387: Data Mining and Concept Learning, Winter 2014

Assignment 3

Due Date: Monday March 24, 2014

Here is the hypothesis I would like you to explore in this paper:

Association rule mining and clustering can help inform the classification process.

To do so, I will ask you to select three data sets of your choice that come in the form of a training set and a testing set. In the first phase, I will ask you to work with the training sets only (of course, you can divide them into training and validation sets to help guide your research) and explore different ways of combining the information you obtain from the association rule mining and clustering processes to improve classification performance. Hopefully, you will be able to design a method that improves upon classification in the three domains that you selected. In a second phase, I will ask you to test your results on the testing set that was put aside during the exploration phase (phase 1). Please compare your results to the results obtained by the base methods (i.e., whatever base classifier you use on the data you have transformed according to association rule mining and clustering, also use on the un modified data)

There certainly has been work on using clustering to improve upon classification (you even read one such paper in Week 8). I am less familiar with the use of association rule mining in this fashion. Feel free to get inspiration from any previous work you read on either subjects, but please cite your sources.

Please note that, this being an assignment rather than a project, it is totally acceptable for you to implement methods previously proposed in the literature. You do not need to design your own novel approach (though you are certainly not discouraged from doing so, either, should you wish to). The main objective of this assignment is to familiarize yourself with other modes of learning (association rule mining and clustering) than classification. Please experiment with different algorithms within these categories.

And, have fun doing the assignment!