

Ensemble Classification Methods

Ian Johnson

Abstract:

Ensemble classification is a method of classification whereby a large set of weakly-trained classifiers functions as one large classifier. Each individual classifier in the set performs a classification on new data, and their classifications are aggregated through some form of voting strategy to decide what the predicted class label should be.

In today's tutorial, we take a cursory look at ensemble classification. We start with a simple introduction of the concept of ensemble classification, and then we look at associative classification as an example. We then explore popular ensemble techniques such as bagging, boosting, and stacking, each with code examples in R. Please prepare an R environment on your laptop to code along with me during class. At the end of the tutorial, attendees should have a basic understanding of ensemble classification, how it's used, and when and why it is effective.