

Tutorial: A survey of ensemble classification

Joseph Kasonde

EMIS 8331 Data mining

Southern Methodist University

Spring 2015

Abstract

An ensemble classifier consists of a set of individually trained classifiers whose predictions are combined for classifying new instances. The performance of an ensemble classifier is in most cases superior to individual classifiers. This tutorial begins by briefly reviewing typical base classifiers that are used to create ensemble classifiers. Statistical and intuitive rational for using an ensemble classifier over any base classifier is provided. An overview of the properties and building blocks of ensemble classifiers is presented together with a review of some of the popular ensemble methods such as Boosting and Bagging. Finally diversity generation is presented along with the different combining methods of base classifiers into an ensemble classifier. An example of the ensemble approach is demonstrated using R.

Key word/phrases: base classifier, ensemble classifier, combining methods, diversity generation, rational for ensembles,