Reducing the Error-propagation Effect Associated with Stacking Classifiers

Abdulaziz Alali

November 16, 2015

Abstract

Multi-label classification targets domains characterized by examples that may belong to more than one category at the same time. A common way to address such problems is to induce a separate classifier for each class. Thus, each classifier determines whether its respective category is relevant for a new example or not. By targeting each class independently, this technique, known as *Binary Relevance (BR)*, assumes that classes are independent of each others, which may not necessarily be the case in all domains. For example, an image of a 'Beach' scene will likely be tagged with the concept 'Ocean' as well. Conversely, the same image is unlikely to represent the topic 'Industry'. To incorporate such class correlations in the *BR* framework, researchers suggest using the classes an example is already known to belong to as additional inputs. Since this information is typically unknown for a previously unseen example, several approaches fill in these values using the outputs of independent classifiers such as those in the *BR* framework. In real-world scenarios, these outputs are prone to errors. Consequently, using them as inputs may reduce the classifiers accuracy. The presented work suggests two ways to address this so-called *error-propagation* problem. The first method reduces the dependency on the error-prone inputs by eliminating weak class dependencies from the final model. The second proposed solution is to compare the probabilistic classification confidences of the independent models with their dependent counterparts, and then choose the more confident classification. Experiments on a broad set of benchmark datasets indicate that a combination of the two approaches yields a boost in classification accuracy when compared to using the dependent models alone.