**Multiclass Classification Using SVM**

In its most simple type, SVM doesn’t support multiclass classification natively. It supports binary classification and separating data points into two classes. For multiclass classification, the same principle is utilized after breaking down the multiclassification problem into multiple binary classification problems.

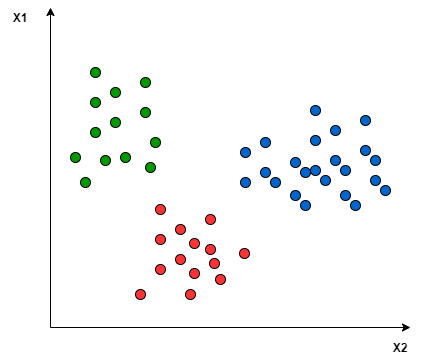
The idea is to map data points to high dimensional space to gain mutual linear separation between every two classes. **This is called a** [***One-to-One* approach**](https://machinelearningmastery.com/one-vs-rest-and-one-vs-one-for-multi-class-classification/)**, which breaks down the multiclass problem into multiple binary classification problems. A binary classifier per each pair of classes.**

**Another approach one can use is** [***One-to-Rest***](https://machinelearningmastery.com/one-vs-rest-and-one-vs-one-for-multi-class-classification/)**. In that approach, the breakdown is set to a binary classifier per each class.**

A single SVM does binary classification and can differentiate between two classes. So that, according to the two breakdown approaches, to classify data points from classes data set:

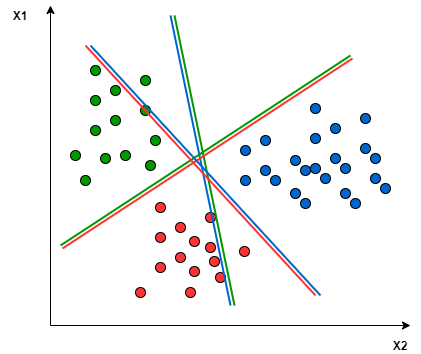
* In the *One-to-Rest* approach, the classifier can use  SVMs. Each SVM would predict membership in one of the  classes.
* In the *One-to-One* approach, the classifier can use SVMs.

Let’s take an example of 3 classes classification problem; green, red, and blue, as the following image:

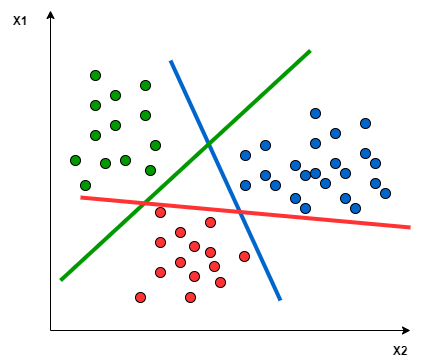


Applying the two approaches to this data set results in the followings:

In the *One-to-One* approach, we need a hyperplane to separate between every two classes, neglecting the points of the third class. This means the separation takes into account only the points of the two classes in the current split. For example, the red-blue line tries to maximize the separation only between blue and red points. It has nothing to do with green points:



In the *One-to-Rest* approach, we need a hyperplane to separate between a class and all others at once. This means the separation takes all points into account, dividing them into two groups; a group for the class points and a group for all other points. For example, the green line tries to maximize the separation between green points and all other points at once:



One of the most common real-world problems for multiclass classification using SVM is text classification. For example, classifying news articles, tweets, or scientific papers.