Multiclass Classification

Multi-class Classification

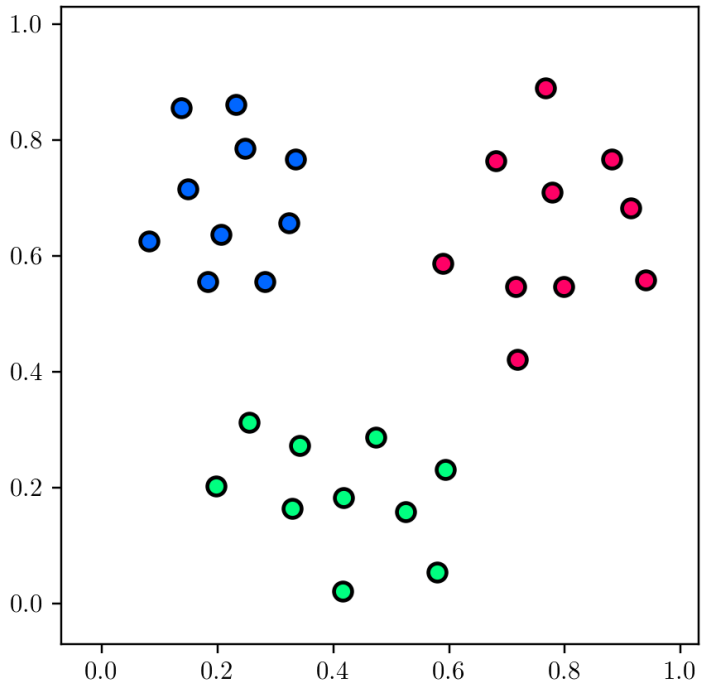
- Let $\mathcal{X} = \{(x_i, y_i) \mid x_i \in \mathbb{R}^n, y_i \in \{1, \ldots, k\}, i \leq m\}$ be a labeled set of points with more than two classes
- In order to define a predictor $f : \mathcal{X} \to \{1, \ldots, k\}$ we can use to strategies:
  - Define a specific Machine Learning algorithm whose prediction function has the set of all the labels as range
  - Generalize a binary classifier to the case with more than 2 labels
- We will focus on the second strategy and consider two approaches: *1-vs-all* and *1-vs-1*

Multi-class classifiers 1-vs-all

- For the 1-vs-all strategy the idea is to build k binary classifiers and combine them to obtain a prediction function for the whole label set $Y = \{1, \ldots, k\}$

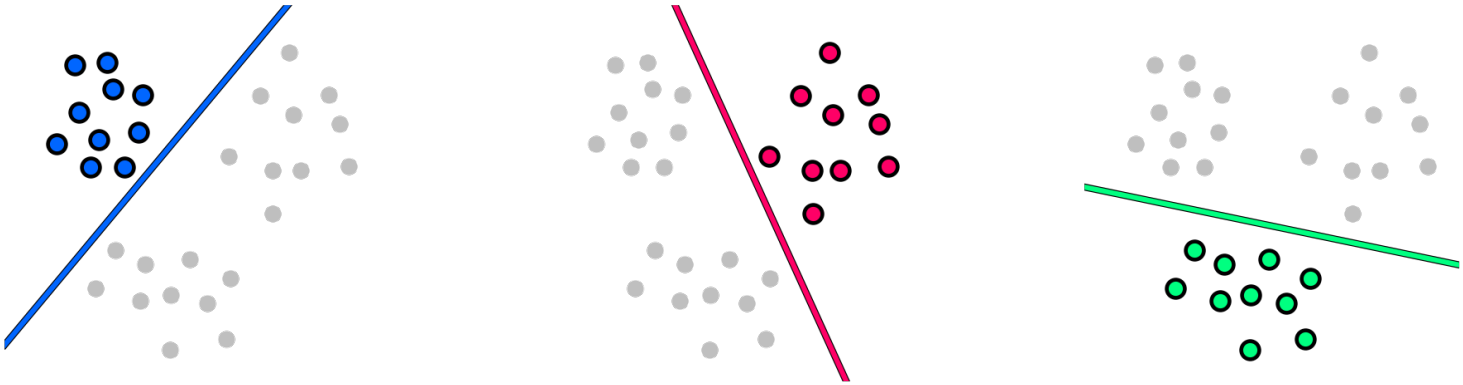
Multi-class classifiers 1-vs-all

- For the 1-vs-all strategy the idea is to build k binary classifiers and combine them to obtain a prediction function for the whole label set $Y = \{1, \ldots, k\}$
- Consider the following picture:



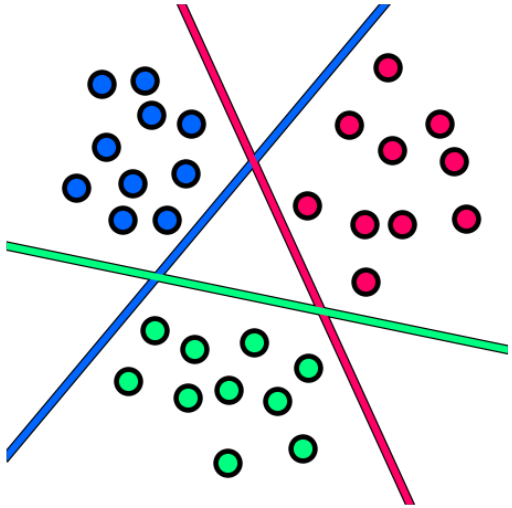
Multi-class classifiers 1-vs-all

- For every class $j \in Y$ we can build a binary classifier that separates the element in the $j - th$ class from all the other elements



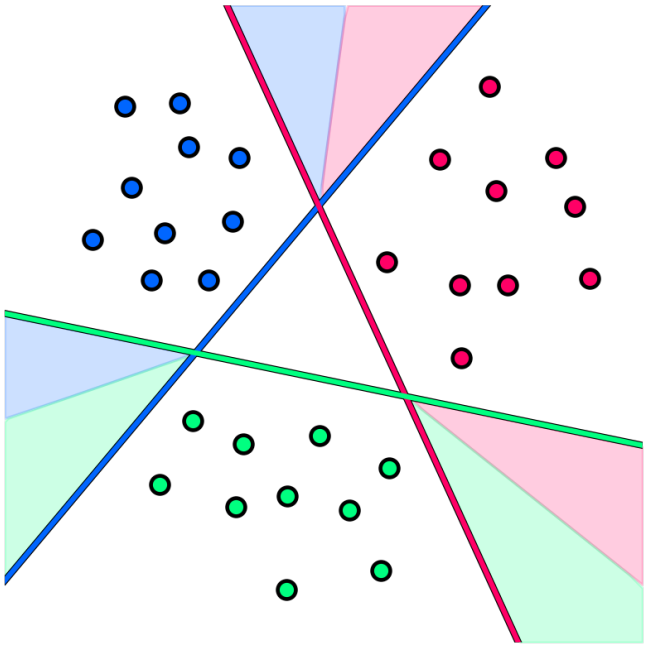
Multi-class classifiers 1-vs-all

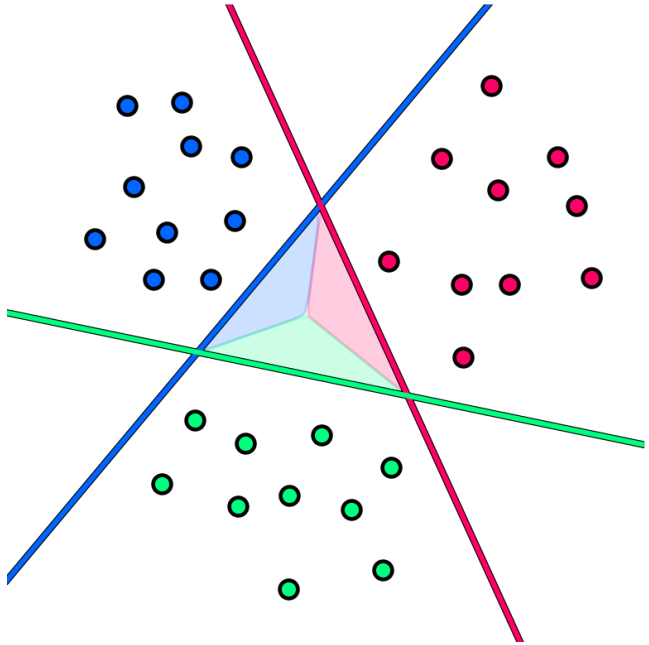
- Once all the k classifiers are build, we can divide the space into k distinct regions



Multi-class classifiers 1-vs-all

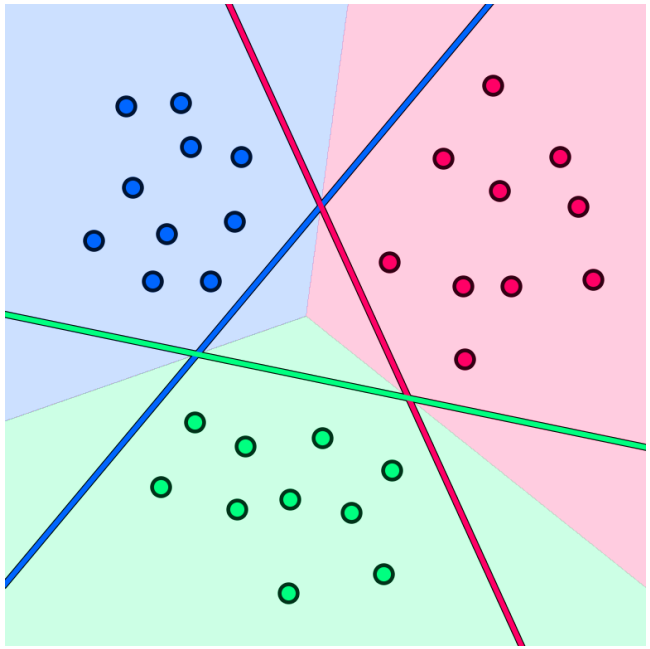
- In this case, in order to classify the points in the common regions, we can use the distance of the points from the separating lines





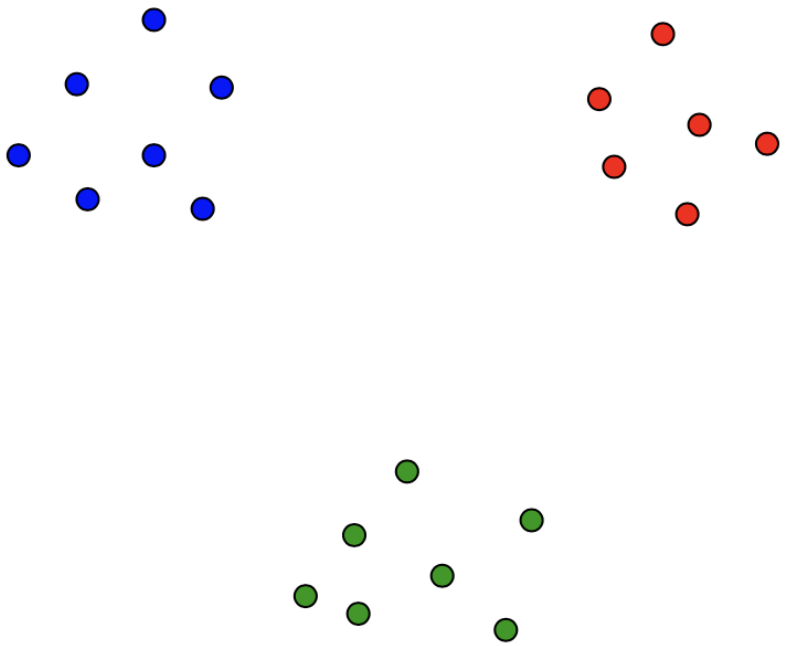
Multi-class classifiers 1-vs-all

- Putting all together we obtain a classifier able to predict the label of any point in the ambient space



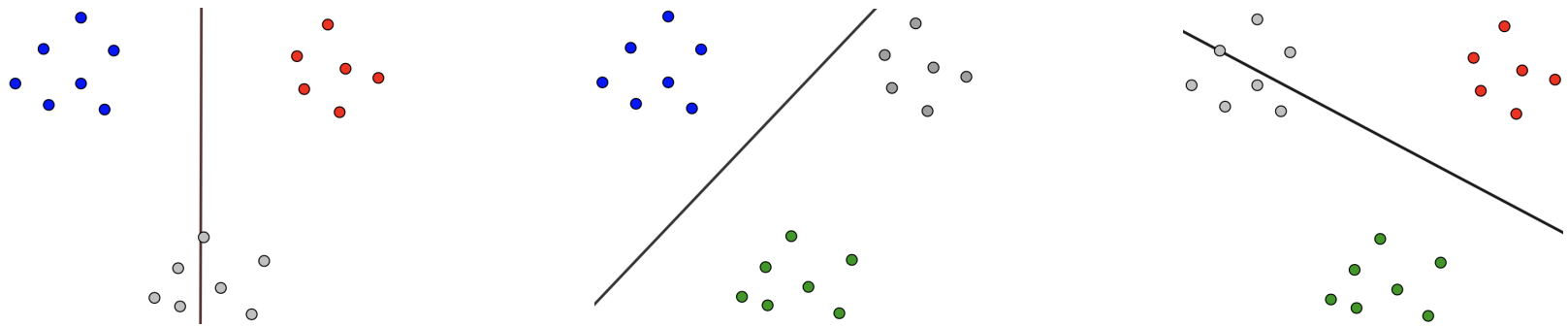
Multi-class classifiers 1-vs-1

- For the 1-vs-1 strategy we build a classifier for every pair of classes. If Y has k different classes we build $\frac{k(k-1)}{2}$ different classifiers
- Consider the following picture:



Multi-class classifiers 1-vs-1

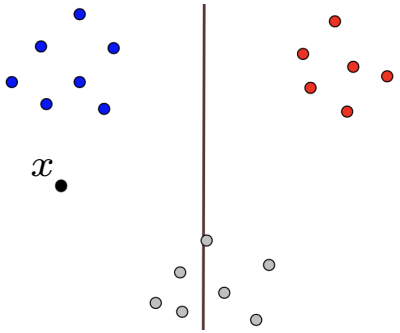
- For every pair of different classes $i, j \in Y$ we build a binary classifier $f_{i,j}(x) : \mathcal{X} \to \{i, j\}$:



Classificatori multi-classe 1-vs-1

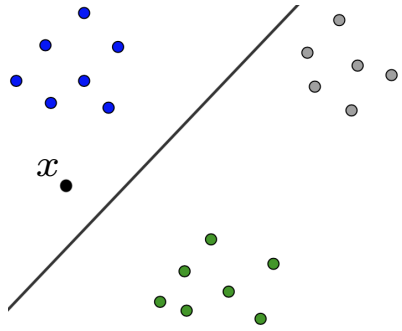
- Once we have all the classifiers f_{ij} we implement a voting system to assign the labels:

$f_{1,2}(x) = \text{Blu}$



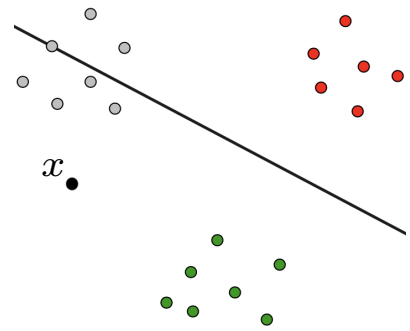
Classes	Votes
Blue	1
Red	0
Green	0

$f_{1,3}(x) = \text{Blu}$



Classes	Votes
Blue	2
Red	0
Green	0

$f_{2,3}(x) = \text{Verde}$



Classes	Votes
Blue	2
Red	0
Green	1

The multi-class classifier $f : \mathcal{X} \to Y$ assign to the element x the label that got the most votes:

$$f(x) = \text{Blue}$$