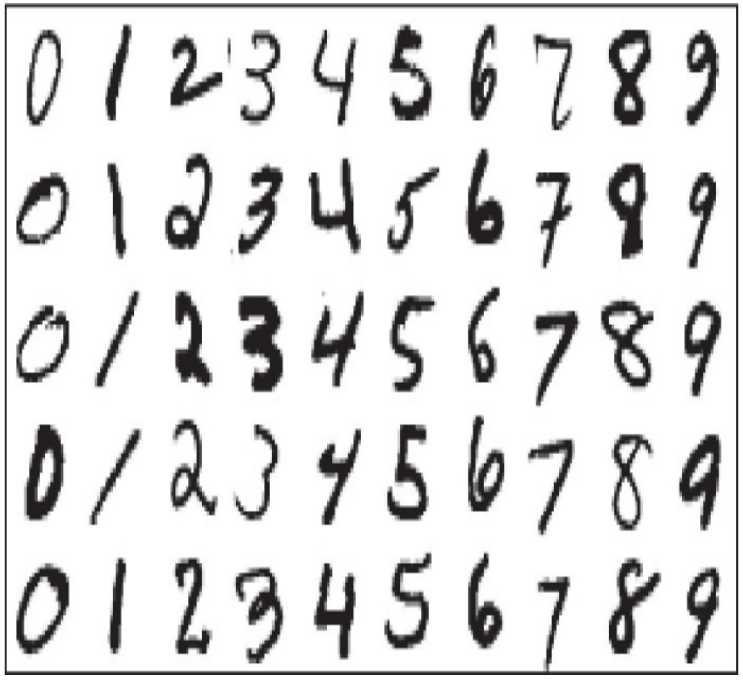
Supervised learning

Lectures Scheduling

- Practical session with Python 3 and scikit-learn
- Website : `http://scikit-learn.org/stable/install.html`

Some questions

Some examples of classification problems :



IRIS dataset

Iris Versicolor

Iris Setosa

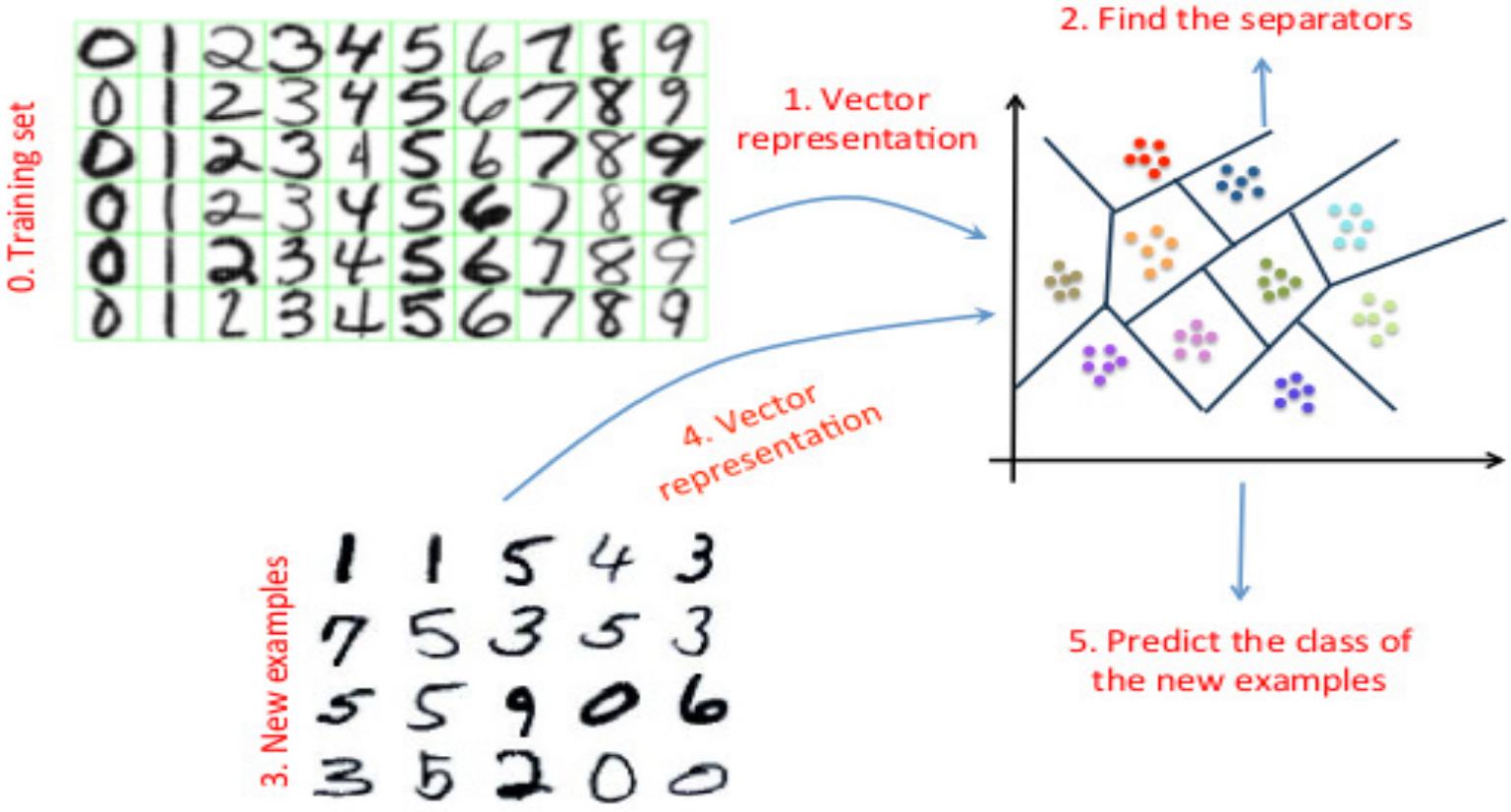
Iris Virginica

The approach

- Find a relevant representation of the data? Need of an expert knowledge on data
- Predict automatically the label of a new observation?

Iris Versicolor **Iris Setosa** **Iris Virginica**

The approach



Context

A two–stage procedure :

- Represent each observation i by a vector $x_i \in \mathcal{X} \subset \mathbb{R}^d$
- $\mathcal{X}$: espace of (features).
- Goal : define a function f mapping each observation x_i on its label $f(x_i) \in \mathcal{Y}$
- Classification : $\mathcal{Y} = \{1, \cdots, K\}$. The function f is called a classifier
- Binary classification $|\mathcal{Y}| = 2$. For example $\mathcal{Y} = \{-1, 1\}$