Unsupervised Learning Laboratories

Alex Rodriguez, Ilaria Vascotto

Fall Semester 2024

1 Data Generation and Visualization

These random datasets will be useful in later laboratories to test some of the algorithms seen in class.

Try to solve the following exercises using only the `numpy` and `matplotlib` python libraries.

1.1 Exercise 1

- Generate a dataset drawn from two independent Gaussians, each of which represents a class. You are free to select the number of desired datapoints, the mean and variance of the two distributions.

- Flip randomly a percentage of the labels to introduce some noise and plot the results.

Given the binary classification setting we are considering, a flip implies that a point in class 0 will be assigned to class 1 and viceversa.

1.2 Exercise 2

The Swiss Roll is defined by the following mapping:

$$x = \phi \cos(\phi) \ , \ y = \phi \sin(\phi) \ , \ z = \psi$$

with $\phi \in (1.5\pi, 4.5\pi)$ and $\psi \in (0, 10)$.

Create a dataset of $n = 1000$ points and plot them both in 2d and 3d.

1.3 Exercise 3

Find the equation describing the Klein bottle problem. Generate a dataset and plot it as in Exercise 2.