Machine Learning

Support Vector Machines

Fabio Vandin

November 25^{th}, 2022

Classification and Margin

Consider a classification problem with two classes:

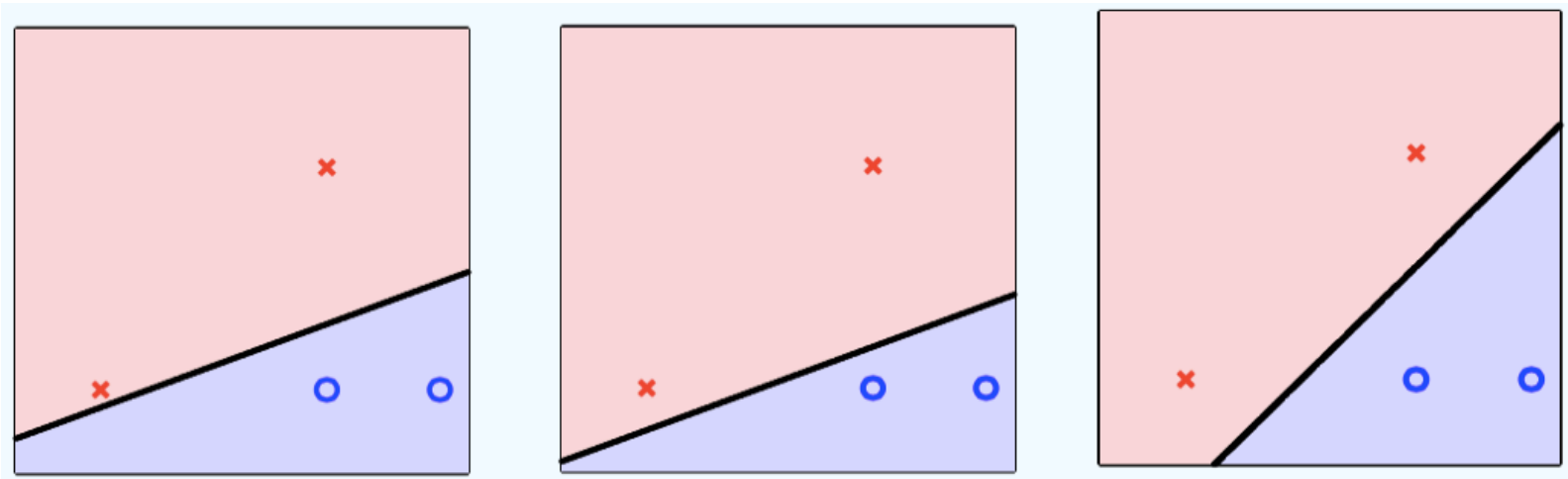
- instance set $\mathcal{X} = \mathbb{R}^d$
- label set $\mathcal{Y} = \{-1, 1\}$.

Training data: $S = ((\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m))$

Hypothesis set $\mathcal{H}$ = halfspaces

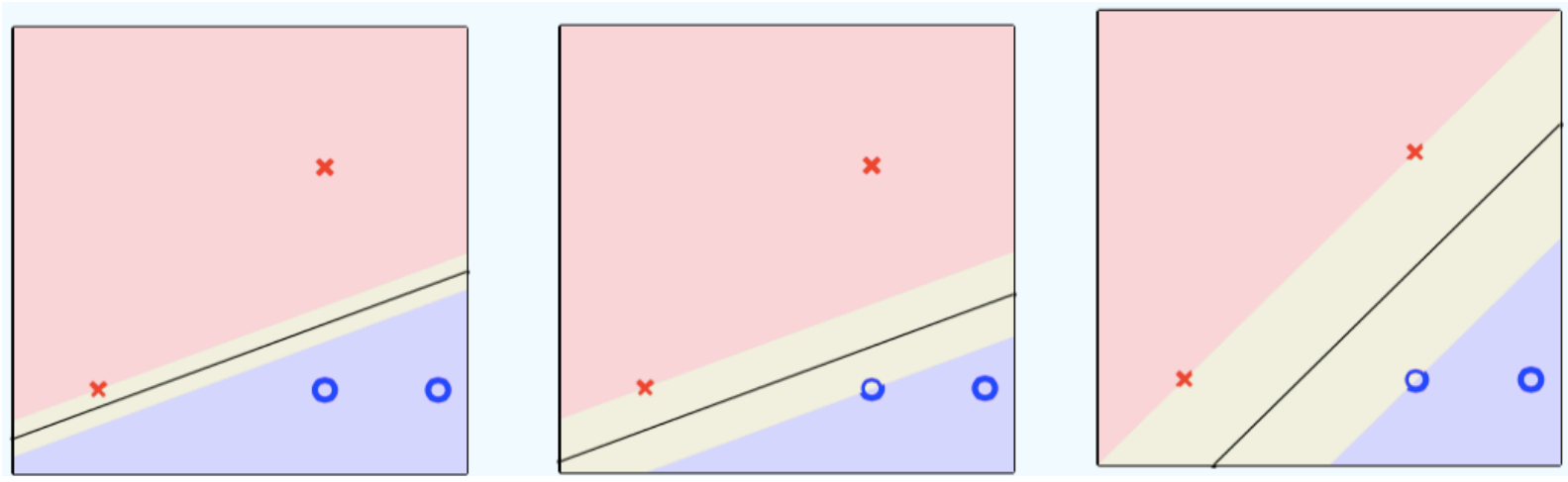
Assumption: data is linearly separable $\Rightarrow$ there exist a halfspace that perfectly classify the training set

In general: multiple separating hyperplanes: $\Rightarrow$ which one is the best choice?



Classification and Margin

The last one seems the best choice, since it can tolerate more "noise".



Informally, for a given separating halfspace we define its *margin* as its minimum distance to an example in the training set S.

Intuition: best separating hyperplane is the one with largest margin.

How do we find it?