Perceptron Algorithm

Given:

- reference set $X = \{\vec{x}_1, \vec{x}_2, \ldots \vec{x}_m\}$

1. Initialise $\vec{w} := 0$ and $t := 0$.
2. Repeat:
   - For all $\vec{x}_i \in X$:
     (a) Classify $\vec{x}_i$ by $\vec{w} \cdot \vec{x}_i \geq t$.
     (b) if correct then continue, else:
        – if false negative, $t := t - 1$, $\vec{w} := \vec{w} + \vec{x}_i$
        – if false positive, $t := t + 1$, $\vec{w} := \vec{w} - \vec{x}_i$
   - until all X correctly classified.

Perceptron Algorithm

Advantages:

- guaranteed convergence when linearly separable
- very fast on test data

Disadvantages:

- thrashes when not linearly separable

The perceptron algorithm is a gradient descent method, but doesn't get stuck in local maxima — either converges to global optimum or never converges.