Lectures

Ilia Kamyshev

January 7, 2021

Contents

1 Rosenblatt's perceptron

1.1 Introduction

Recall: for an input $x \in \mathbb{R}^n$, the parametrized function describing the mapping computed by a perceptron is:

$$F(\omega, x) = F((\omega_1, \omega_2, \ldots, \omega_n, \theta), (x_1, x_2, \ldots, x_n)) = sgn(\sum_{i=1}^{n} \omega_i x_i - \theta)$$

Or in the vector form:

$$F(\boldsymbol{\omega}, \boldsymbol{x}) = sgn(\boldsymbol{\omega^T x})$$

This perceptron can only classify patterns which are *linearly seprable.*

Theorem 1. *Suppose $C_1 \bigcup C_2 = C$ are linearly separable classes over the training set $z \in Z^T$ with the assumption $z_t = \{\boldsymbol{x_t}, y_t \in C_1\}$, and perceptron's response r_t with mistake $e_t = r_t - y_t \neq 0$ can be corrected by applying the learning rule to its current state $\boldsymbol{\omega} \in \boldsymbol{\Omega}^T$:*

$$\boldsymbol{\omega_{t+1}} = \boldsymbol{\omega_t} + e_t \boldsymbol{x_t}$$

Then perceptron's error correction algorithm converges in k number of steps with following assumptions: training input is bounded by Euclidean norm $\|\boldsymbol{x_t}\| \leq R$ and $\boldsymbol{\omega_k^T x_t} \geq \gamma$ for $t = 1..T$, where $\gamma > 0$. Initial state $\boldsymbol{\omega_0} = 0$. Note, that γ uses to be sure that some example is classified correctly.

Proof: Expand equation for learning rule at t steps as:

$$\boldsymbol{\omega_{t+1}} = \boldsymbol{x_1} + \boldsymbol{x_2} + \ldots + \boldsymbol{x_t}$$

Multiply each side by $\boldsymbol{\omega_k^T}$:

$$\boldsymbol{\omega_k^T}\boldsymbol{\omega_{t+1}} = \boldsymbol{\omega_k^T}\boldsymbol{x_1} + \boldsymbol{\omega_k^T}\boldsymbol{x_2} + \ldots + \boldsymbol{\omega_k^T}\boldsymbol{x_t}$$

From the assumption $\boldsymbol{\omega_k^T}\boldsymbol{x_t} \geq \gamma$:

$$\boldsymbol{\omega_k^T}\boldsymbol{\omega_{t+1}} \geq k\gamma$$

From *Cauchy-Schwarz inequality*:

$$\left|\boldsymbol{\omega_k^T}\boldsymbol{\omega_{t+1}}\right| \leq \left\|\boldsymbol{\omega_k^T}\right\| \cdot \left\|\boldsymbol{\omega_{t+1}}\right\|$$