

Персептрон

Определение

В 1958 году американский физиолог Ф. Розенблatt, изучая феномен восприятия, предложил математическую модель нейрона, которая получила название *персептрон*:

$$\mathbf{u} \mapsto \text{sign}(\langle \mathbf{w}, \mathbf{u} \rangle) \quad (\mathbf{w}, \mathbf{u} \in \mathbb{R}^m).$$

Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review*, 1958, vol. 65, pp. 386-407.

Персепtron

Алгоритм Розенблатта

Вход: Набор примеров вида

$$(\mathbf{u}_1, y_1), \dots, (\mathbf{u}_n, y_n).$$

Основной блок:

$$\mathbf{w}_0 := (0, \dots, 0).$$

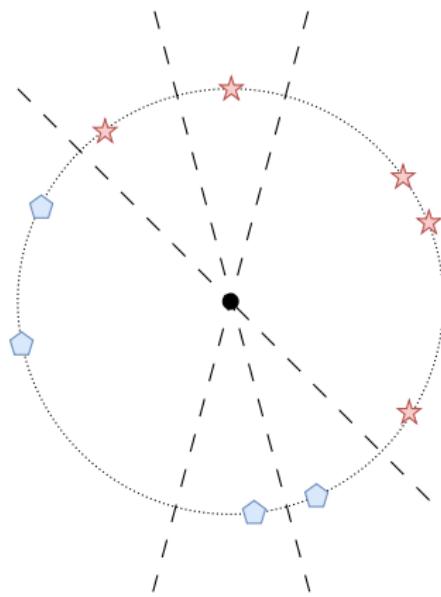
Цикл ($t = 0, 1, \dots$):

Если: существует индекс i_t
($1 \leq i_t \leq n$) такой, что
 $\langle \mathbf{w}_t, y_{i_t} \mathbf{u}_{i_t} \rangle \leq 0$, **то:**

$$\mathbf{w}_{t+1} := \mathbf{w}_t + y_{i_t} \mathbf{u}_{i_t};$$

иначе: выход из цикла.

Выход: \mathbf{w}_t .



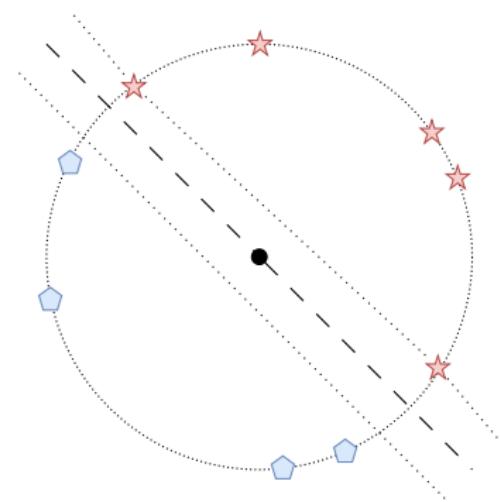
Персепtron

Теорема Новикова

Зазор (margin):

$$\rho_* := \max_{\|\mathbf{w}\|_2=1} \rho(\mathbf{w}),$$

$$\rho(\mathbf{w}) := \min_{i=1,\dots,n} \langle \mathbf{w}, y_i \mathbf{u}_i \rangle.$$



Персептрон

Теорема Новикова

Теорема (Новикова)

Предположим, что для входного набора примеров выполняется условие $\rho_* > 0$. Тогда алгоритм за t шагов, где

$$t \leq \frac{D^2}{\rho_*^2}, \quad D := \max_{i=1,\dots,n} \|\mathbf{u}_i\|_2. \quad (1)$$

построит разделяющую гиперплоскость.

Novikoff A.B. On convergence proofs on perceptrons, *Symposium on the Mathematical Theory of Automata*, 1958, vol. 12, pp. 615-622.

Персепtron

Выводы

Выводы:

- если $m > n$, то данные разделимы;
- если известна величина зазора ρ_* , то имеется оценка для времени обучения;
- если данные не разделимы, то мы получим NP -трудную задачу.