

Лабораторна робота №3

Одношаровий персептрон

Мета: отримати навички розв'язання практичних задач за допомогою одношарового персептрона.

Короткі теоретичні положення

Модель персептрона

Модель персептрона має вигляд, показаний на рис. 3.1.

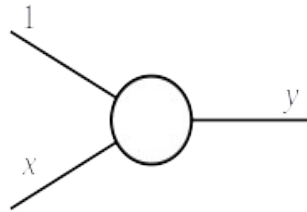


Рис. 3.1. – Модель персептрона

При цьому

$$x \in R^d, \text{ або } x \in \{-1, 1\}^d,$$

$$y \in R, \text{ або } y \in \{-1, 1\}.$$

Таку модель використовують для розв'язання задачі класифікації для двох класів і є ідентичною до задачі

$$y \in \{0, 1\}.$$

Будемо розглядати випадок

$$x \in R^d, y \in \{-1, 1\}.$$

Функціонування персептрона описується наступною залежністю:

$$y = \text{sign}(W^T x - \tau) = f(x, W). \quad (3.1)$$

де τ – деякий поріг, W – вектор вагових коефіцієнтів персептрона.

У геометричній інтерпретації рівняння (3.1) визначає два підпростори

$$\{x : y = 1\} \Leftrightarrow H^+ = \{x : W^T x \geq \tau\}, \quad (3.2)$$

$$\{x : y = -1\} \Leftrightarrow H^- = \{x : W^T x < \tau\},$$

з роздільною гіперплощиною (афінний підпростір розмірності $d - 1$):

$$H = \{x : W^T x - \tau = 0\}. \quad (3.3)$$

Збільшуючи розмірність простору, отримаємо

$$x \in R^d \Rightarrow \tilde{x} \in R^{d+1}, \quad (3.4)$$

де $\tilde{x}_i = x_i, \tilde{x}_{d+1} = 1, i \leq d$,

$$W \in R^d \Rightarrow \tilde{W} \in R^{d+1}, \quad (3.5)$$

де $\tilde{W}_i = W_i, \tilde{W}_{d+1} = \tau, i \leq d$.

Враховуючи (3.4) та (3.5), можна записати

$$W^T x - \tau = \tilde{W}^T \tilde{x}.$$

Навчання перцептрона (алгоритм Розенблатта)

Навчання перцептрона представляє собою процес налаштування вагових коефіцієнтів W . При навчанні нейронної мережі, як правило, математичні вирази для розділяючих поверхонь відсутні. Тому навчання виконується тільки на навчальній вибірці.

Навчальна вибірка (скінчена) задається множиною, що складається з пар вхід-вихід:

$$T = \{(x_1, t_1), \dots, (x_n, t_n)\} = \{(x_i, t_i), i = 1, \dots, n\}, \quad (3.6)$$

де $t_i \in \{-1, 1\}$.

Мета навчання – налаштувати вагові коефіцієнти W таким чином, щоб для будь-яких виконувалось $x^* \in R^d, x^* \notin T$ виконувалось $y = t^*$.

Алгоритм навчання перцептрона Розенблатта:

Даний алгоритм коректно працює лише в тих випадках, коли класи є лінійно розділними.

1. Формуємо множину

$$\tilde{F} = \tilde{S}^+ \cup \left[-\tilde{S}^- \right] \subset R^{d+1},$$

де

$$\begin{aligned} \tilde{S}^+ &= \{x: \text{якщо існує } i \text{ таке, що } t_i = 1, x = x_i\}, \\ \tilde{S}^- &= \{x: \text{якщо існує } i \text{ таке, що } t_i = -1, x = x_i\} \\ -\tilde{S}^- &= \{\tilde{z} : \forall \tilde{x}_i \in \tilde{S}^-, \tilde{z}_i = t_i \tilde{x}_i\}, \end{aligned}$$

і систему

$$\tilde{W}^T \tilde{z} > 0 \text{ для будь яких } \tilde{z} \in \tilde{F}.$$

2. Початок. Вибираємо деякий елемент $\tilde{z} \in \tilde{F}$ як початкове наближення для \tilde{W} . Сформуємо випадкову послідовність (циклічну, у якій елементи з'являються з невизначеною частотою) з елементів \tilde{W} .

3. Тест. Вибираємо випадкове значення $\tilde{z}_{ij} = \text{rand}(\tilde{F})$. Якщо $\tilde{W}^T \tilde{z}_{ij} > 0$, переходимо до п. 3, інакше — до п. 4.

4. Модифікація вагових коефіцієнтів.

$$\tilde{W} = \tilde{W} + \varphi \tilde{z}_{ij},$$

де $\varphi = 1$.

(Операції 4 обумовлені пошуком розв'язку \tilde{W} у формі

$$\tilde{W} = \sum_j \alpha_j \tilde{z}_j, \alpha_j > 0.$$

Крім того

$$\tilde{W}_j \cdot \tilde{z}_j = \tilde{W}_{j-1} \cdot \tilde{z}_{ij} + \varphi > \tilde{W}_{j-1} \cdot \tilde{z}_{ij}.$$

Значення \tilde{W}_j – збільшується, щоб після поточного негативного значення на наступному кроці було отримане додатне (п. 4 виконується тільки у випадку негативного добутку). Переходимо до п. 3.

5. Завершення. Процес навчання закінчується тоді, коли умова $\tilde{W}^T \tilde{z}_{ij} > 0$ буде виконуватися для всіх векторів навчальної вибірки.

Порядок виконання роботи

1. Реалізувати одношаровий персептрон.
2. За допомогою реалізованого персептрона розв'язати задачу. Для цього необхідно випадковим чином сформулювати навчальну та тестову вибірки (у співвідношенні 4:1). Навчити нейронну мережу на навчальній вибірці, використовуючи алгоритм Розенблатта.
3. Перевірити роботу персептрона на тестових даних.
4. Результати роботи оформити звітом, який має містити: постановку задачі, навчальну вибірку даних та їх представлення у графічному виді на R^2 , результати роботи на тестовій множині даних, параметри персептрона, що навчився, вихідний код програми.