Лабораторна робота №2А

Одношаровий персептрон

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

1.1. Теоретичні відомості

Модель персептрона

Модель персептрона має вигляд, показаний на рис. 1.1.

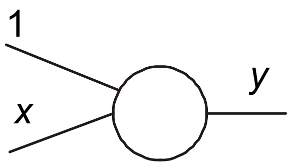

Рис. 1.1. Модель персепртрона

При цьому

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

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

Будемо розглядати випадок

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

Функціонування персептрона описується наступною залежністю:

$$y = \text{sign}(W^{\mathrm{T}} x - \tau) = \eta(x, W), \tag{1.1}$$

[1] *Задача класифікації на два класи. Також може бути* $y \in \{0,1\}$.

де τ — деякий поріг;

W — вектор вагових коефіцієнтів персептрона.

У геометричній інтерпретації рівняння (1.1) визначає два *підпростори*

$$\{x: y=1\} \quad \Leftrightarrow \quad H^{+}=\{x: W^{\mathrm{T}} x \geq \tau\}, \tag{1.2}$$

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

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

$$H=\{x: W^{\mathrm{T}} x-\tau=0\}. \tag{1.3}$$

Збільшуючи розмірність простору, отримаємо

$$x \in R^{d} \Rightarrow \widetilde{x} \in R^{d+1}, \tag{1.4}$$

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

$$W \in R^{d} \Rightarrow \widetilde{W} \in R^{d+1}, \tag{1.5}$$

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

Враховуючи (1.4) та (1.5), можна записати

$$W^{\mathrm{T}} x-\tau=\widetilde{W}^{\mathrm{T}} \widetilde{x},$$

де $\widetilde{x}_{d+1}$ — Bias-нейрон.

Навчання персептрона (алгоритм Розенблатта)

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

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

$$T = \{(x_1, t_1), \ldots, (x_n, t_n)\} = \{(x_i, t_i),\ i = \overline{1, n}\ \}, \tag{1.6}$$

де $t_i \in \{-1, 1\}$.

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

Алгоритм навчання персептрона Розенблатта[2]:

1. Формуємо множину

$$\tilde{F} = \tilde{S}^+ \cup \{-\tilde{S}^-\} \subset R^{d+1}, \text{ де}$$

$$S^+ = \{x : \text{ якщо існує } i \text{ таке, що } t_i = 1, \quad x = x_i\}$$
$$S^- = \{x : \text{ якщо існує } i \text{ таке, що } t_i = -1, \quad x = x_i\}$$

$$-\tilde{S}^- = \{\tilde{z} : \forall \tilde{x}_i \in \tilde{S}^-,\ \tilde{z}_i = t_i \tilde{x}_i\},$$

і систему

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

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 $\tilde{z}_{i_j} = \mathrm{rand}(\tilde{F})$. Якщо $\tilde{W}^T \tilde{z}_{i_j} > 0$, переходимо до п. 3, інакше — до п. 4.

4. *Модифікація вагових коефіцієнтів*. Сформуємо обмежену послідовність

$$\varphi_j : 0 < \underline{\varphi} \le \varphi_j \le \overline{\varphi},$$

$$\tilde{w} = \tilde{w} + \varphi_j \tilde{z}_{i_j}.$$

[2] *Даний алгоритм коректно працює лише в тих випадках, коли класи є лінійно роздільними.*

Переходимо до п. 3.

5. *Завершення*. Процес навчання закінчується тоді, коли умова $\widetilde{W}^{T}\widetilde{z}_{i_j} > 0$ буде виконуватися для всіх векторів навчальної вибірки.[3]

Зауваження.

1. У базовому алгоритмі навчання персептрона $\varphi_k = 1$, але найчастіше вибирають

$$\varphi_k = \frac{1}{\left\| \widetilde{z}_{i_j} \right\|},$$

для нормування множини $\widetilde{F}$ таким чином, щоб усі його вектори мали одиничну довжину.

2. Операції 4 обумовлені пошуком розв'язку $\widetilde{W}$ у формі

$$\widetilde{W} = \sum_j \alpha_j \widetilde{z}_{i_j}, \;\; \alpha_j > 0.$$

Крім того

$$\widetilde{W}_j \cdot \widetilde{z}_{i_j} = \widetilde{W}_{j-1} \cdot \widetilde{z}_{i_j} + \varphi_j \left\| \widetilde{z}_{i_j} \right\|^2 > \widetilde{w}_{j-1} \cdot \widetilde{z}_{i_j}.$$

Значення $\widetilde{W}_j$ — збільшується, щоб після поточного негативного значення на наступному кроці було отримане додатне (п. 4 виконується тільки у випадку негативного добутку).

---

[3] *Теоретично доведено, що якщо класи є лінійно роздільними, алгоритм Розенблатта зійдеться за скінчену кількість кроків.*

1.2. Порядок виконання роботи

1. Реалізувати одношаровий персептрон, використовуючи такі мови програмування як Python, C++.

2. За допомогою реалізованого персептрона розв'язати задачу згідно з номером варіанту. (Номер варіанту визначається за номером у списку групи.) Для цього на основі відповідного файлу (ім'я data*номер_варіанту*.csv) необхідно випадковим чином сформувати навчальну та тестову вибірки (у співвідношенні 70:30). Навчити нейронну мережу на навчальній вибірці, використовуючи алгоритм Розенблатта.

! Для отримання **максимального балу** за лабораторну роботу **заборонено** використовувати бібліотеки з вже реалізованими аналогічними алгоритмами (потрібно реалізовувати алгоритми самостійно). Вбудовані алгоритми використовувати лише для порівняння власно-запрограмованого алгоритму.

3. Перевірити роботу персептрона на тестових даних. Порівняти результати з аналогічними результатами, які отримані в результаті використання вбудованих функцій.

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. Прокоментувати код (що кожна строка робить, окрім import / library).