Лабораторна робота 1

Лінійний бінарний класифікатор

Мета роботи: розробити програму для лінійної бінарної класифікації вхідного набору даних за допомогою нейрона Маккаллока-Піттса і дослідити її роботу для різних наборів векторів лінійно-роздільних і лінійно-нероздільних класів.

Теоретичні відомості

Математична модель штучного нейрона

При побудові штучних нейронних мереж використовується математична модель штучного нейрона, блочна діаграма якого зображена на рис. 1.1. На цій діаграмі можна виділити такі елементи:

1. *Вхідні сигнали* нейрона $x_1, x_2, \dots, x_m$, які можна згрупувати у вхідний вектор $\boldsymbol{x} = \{x_1, x_2, \dots, x_m\}$. Компоненти вхідного вектору є дійсними числами і зазвичай описують *ознаки* (атрибути, характеристики) деякого об'єкту, що розглядається.

2. *Синапси* або *синаптичні зв'язки*, які характеризуються *ваговими коефіцієнтами* (дійсними числами) або просто *вагами* $w_1, w_2, \dots, w_m$. Сигнал

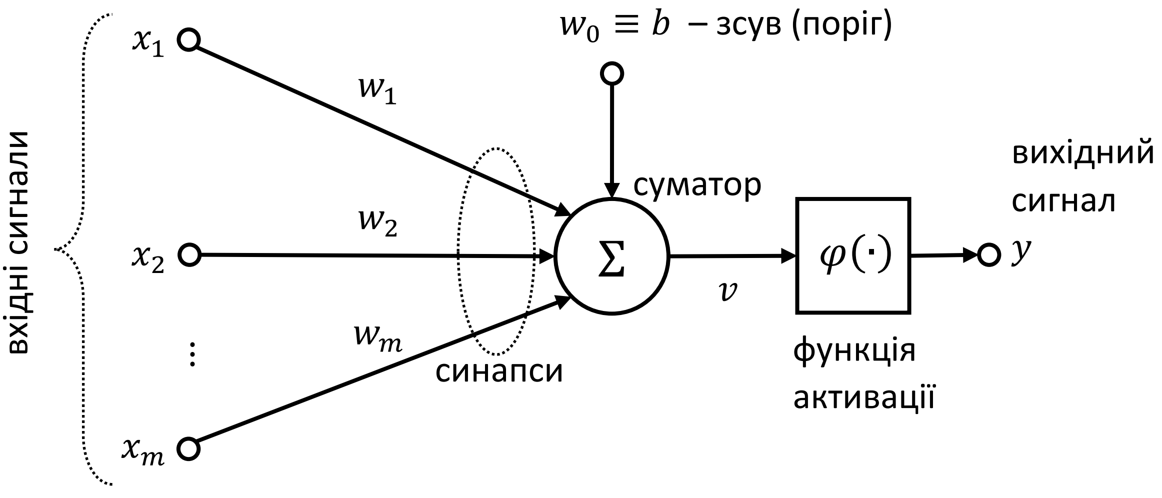


Рис. 1.1. Блочна діаграма штучного нейрона.

x_j на вході синапсу j множиться на вагу w_j. Як і вхідні сигнали, вагові коефіцієнти групують у вектор ваг $\boldsymbol{w} = \{w_1, w_2, \dots, w_m\}$.

3. *Суматор* Σ складає всі зважені вхідні сигнали $w_j x_j$ і ще один коефіцієнт w_0, який називається *порогом* (threshold) або *зсувом* (bias). Цей коефіцієнт іноді позначається символом b. Таким чином, суматор формує на своєму виході сигнал v, яка називається *індукованим локальним полем* або *потенціалом активації* або *постсинаптичним потенціалом*:

$$v = \sum_{j=1}^{m} w_j x_j + w_0 = \boldsymbol{w} \cdot \boldsymbol{x} + w_0. \tag{1.1}$$

4. *Функція активації* φ, яка застосовується до індукованого локального поля v, утворює вихідний сигнал

$$y = \varphi(v) = \varphi(\boldsymbol{w} \cdot \boldsymbol{x} + w_0). \tag{1.2}$$

У сучасних штучних нейронних мережах функція активації може мати різноманітний вигляд і майже завжди є нелінійною.

Нейрон з *пороговою* функцією активації (див. рис. 1.2)

$$\varphi(v) = \begin{cases} 1, \text{якщо } v \geq 0, \\ 0, \text{якщо } v < 0 \end{cases} \tag{1.3}$$

називається *нейроном Маккаллока-Піттса* (McCalloch and Pitts) або *одиночним персептроном*.

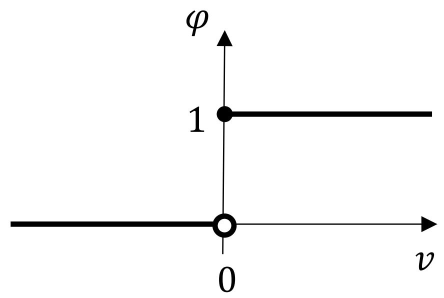


Рис. 1.2. Порогова функція активації

Нейрон Маккаллока-Піттса може бути використаний для вирішення задач *лінійної бінарної класифікації* даних.

Задача лінійної бінарної класифікації

Нехай дана навчальна множина $\boldsymbol{T} = (\boldsymbol{x}_i, d_i)_{i=1}^{N}$, яка складається з N маркованих прикладів, тобто пар вхідних m-вимірних векторів $\boldsymbol{x}_i =$

$x_{i1}, x_{i2}, \dots, x_{im}$ і відповідних їм міток класів d_i, які приймають скалярне значення 0 або 1 в залежності від класу, до якого належить вектор $\boldsymbol{x}_i$.

Задача полягає у знаходженні такої *лінії* (для двовимірного простору векторів) або *гіперплощини* (для загального випадку багатовимірного простору векторів), яка розділяє вхідні вектори на два класи (рис. 1.3). Вочевидь, для вирішення такої задачі класи вхідних векторів мають бути *лінійно-роздільними*.

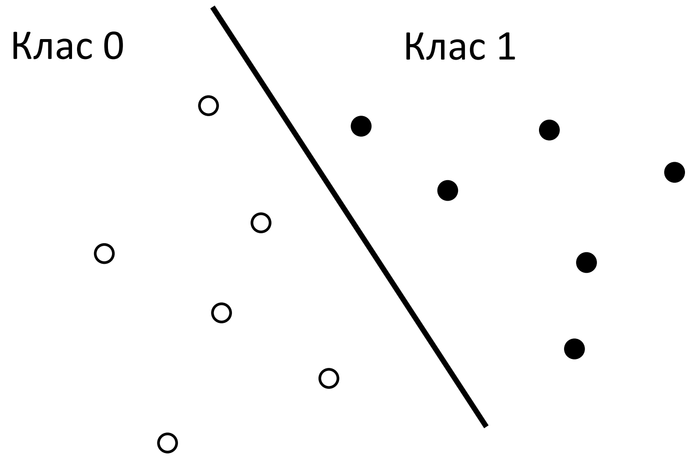


Рис. 1.3. Лінійно-роздільні класи векторів.

У випадку використання нейрона Маккаллока-Піттса для вирішення задачі, що розглядається, слід підібрати вагові коефіцієнти $w_0, w_1, \dots, w_m$ таким чином, щоб для всіх векторів $\boldsymbol{x}_i$ вихідні сигнали нейрона $y_i = \varphi(\boldsymbol{w} \cdot \boldsymbol{x}_i + w_0)$ співпадали з бажаними відгуками d_i. Процес підбору (налаштування, адаптації) вагових коефіцієнтів називається *навчанням* нейрона. Після навчання нейрона рівняння гіперплощини, яка розділяє простір вхідних векторів $\boldsymbol{x}$ на дві області, що відповідають різним класам, матиме вигляд

$$\boldsymbol{w} \cdot \boldsymbol{x} + w_0 = 0 \quad \text{або} \quad w_n x_n + \dots + w_2 x_2 + w_1 x_1 + w_0 = 0. \qquad (1.4)$$

Вхідні вектори $\boldsymbol{x}$, для яких $\boldsymbol{w} \cdot \boldsymbol{x} + w_0 \geq 0$, належать до класу 1, а інші вектори, для яких $\boldsymbol{w} \cdot \boldsymbol{x} + w_0 < 0$, належать до класу 0.

Для двовимірних вхідних векторів $\boldsymbol{x}$ межею між двома класами є пряма лінія:

$$w_2 x_2 + w_1 x_1 + w_0 = 0 \quad \text{або} \quad x_2 = -\frac{w_1}{w_2} x_1 - \frac{w_0}{w_2}. \qquad (1.5)$$

Вхідні вектори, розташовані вище цієї прямої, належать до класу 1, а вектори, розташовані нижче прямої, належать до класу 0.

Алгоритм навчання нейрона Маккаллока-Піттса

Алгоритм навчання нейрона Маккаллока-Піттса як частинного випадку персептрона (одиночного персептрона) був запропонований Ф. Розенблаттом у 1958 році:

1. Ініціалізувати синаптичні ваги і поріг нейрона нульовими (або випадковими) значеннями $w_j = 0,\ j = \overline{0, m}$.
2. Встановити лічильник *епох* $p = 1$ (епоха навчання – цикл пред'явлення нейрону всіх векторів навчальної множини).
3. Встановити лічильник вхідного вектору навчальної вибірки $i = 1$ (початок нової епохи).
4. Для поточних значень ваг w_j і вектору $\boldsymbol{x}_i$ розрахувати вихідний сигнал $y_i = \varphi(\boldsymbol{w} \cdot \boldsymbol{x}_i + w_0)$, де $\varphi(v)$ – порогова функція активації.
5. Значення ваг нейрона скорегувати за *дельта-правилом Відроу-Хоффа для порогової функції активації*:

$$\begin{aligned} w_0 &\coloneqq w_0 + \eta\, e_i, \\ w_j &\coloneqq w_j + \eta\, e_i\, x_{ij},\ \ j = \overline{1, m}, \end{aligned} \tag{1.6}$$

   де $e_i = d_i - y_i$ – *сигнал помилки* для поточного вектору $\boldsymbol{x}_i$, η – коефіцієнт *швидкості навчання* (learning rate), зазвичай обирається в межах $0 < \eta \leq 1$. Якщо початкові вагові коефіцієнти ініціалізовані нульовими значеннями (пункт 1), то для пошуку рівняння гіперплощини (4), яка розділяє вхідні вектори на два класи, вибір значення коефіцієнту швидкості навчання η неважливий, адже коефіцієнти рівняння (4) визначені з точністю до множника. Таким чином значення η можна взяти рівним 1.
6. Збільшити лічильник вхідного вектору на одиницю ($i \coloneqq i + 1$). Якщо $i \leq N$, то перейти до пункту 4, інакше – епоха навчання завершується.
7. Якщо критерії зупинки алгоритму навчання не виконуються, то збільшити лічильник епох на одиницю ($p \coloneqq p + 1$) і перейти до наступної епохи, тобто до пункту 3.

Відповідно до *теореми про збіжність персептрона*, доведеної Ф. Розенблаттом, незалежно від початкових значень вагових коефіцієнтів для навчальних векторів *лінійно-роздільних* класів наведений алгоритм навчання збігається, тобто вагові коефіцієнти, починаючи з деякої епохи, перестають змінюватись, адже нейрон правильно класифікує всі вектори і сигнал помилки e_i дорівнює 0 для всіх $i = \overline{1, N}$. У такому разі відсутність помилок класифікації може стати критерієм зупинки алгоритму.

Якщо такий алгоритм застосувати до векторів *лінійно-нероздільних* класів, то процес корекції ваг стане нескінченим, а алгоритм буде прагнути мінімізувати помилки навчання.

Враховуючи, що лінійна роздільність класів вхідних векторів не може бути завчасно гарантованою, практичними критеріями зупинки алгоритму (пункт 7) можуть бути виконання однієї або декількох умов:

1) досягнення достатньої точності навчання, коли кількість невірно класифікованих векторів не перевищує задану кількість (враховуючи, що величини d_i і y_i можуть приймати лише значення 0 або 1, кількість невірно класифікованих векторів дорівнює сумарній абсолютній помилці $E = \sum_{i=1}^{N} |d_i - y_i|$);
2) досягнення заданої максимальної кількості епох;
3) точність навчання перестає покращуватись протягом заданої кількості епох;
4) досягнення заданого максимального часу обрахунків.

Зверніть увагу, що сумарну абсолютну помилку (або кількість невірно класифікованих векторів) E слід розраховувати *після* завершення кожної епохи, а *не під час* її виконання.

Сумарна абсолютна помилка E може *немонотонно* зменшуватись з кількістю пройдених епох, тому після зупинки алгоритму внаслідок виконання умов 2-4 як результат навчання нейрона обирають не кінцеві значення ваг і порогу, а оптимальні значення – ті, що відповідали найменшій помилці E протягом виконання алгоритму. Для цього після проходження чергової епохи при кожному зменшенні мінімальної сумарної помилки E слід передбачити збереження (оновлення) оптимальних значень ваг і порогу.

Література для поглибленого вивчення теоретичного матеріалу

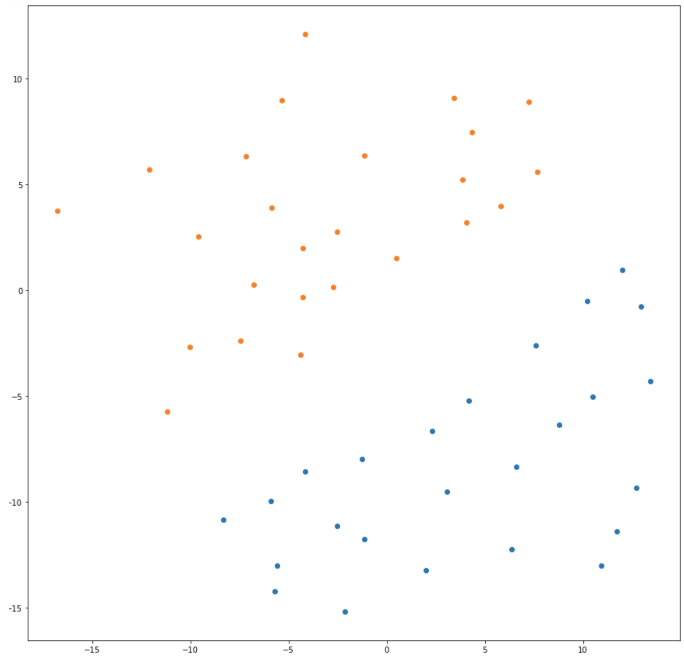
1. Хайкин С. Нейронные сети: полный курс, 2-е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с. (стор. 42 – 48, 91 – 93).

Завдання для самостійної підготовки до лабораторної роботи

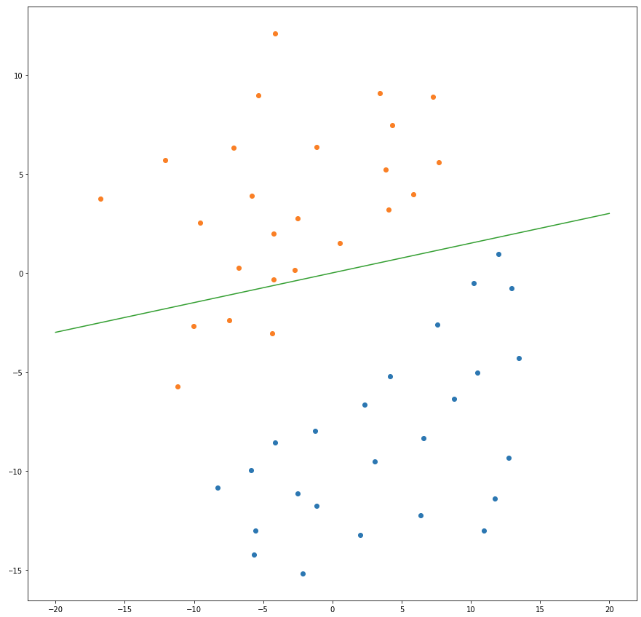
1. Розібратися як працює сервіс Colab (guide).

   Завантажити, відповідно до свого номеру в списку групи, файл до своєї віртуальної машини (наприклад, для номеру 3 sample003.csv). Прочитати дані із файлу за допомогою функції genfromtxt бібліотеки numpy або функції read_csv бібліотеки pandas.

   2. Побудувати зчитані точки на графік. Точки що відносяться до різних класів промалювати різними кольорами. Можна скористатися бібліотекою matplotlib, функцією pyplot.plot. Приклад рисунка:



3. Побудувати графік відповідно завдання нижче. Приклад рисунка:



4. Зберегти файл в форматі <Прізвище>.ipynb

Завдання по варіантах (номер завдання відповідає номеру в списку групи)

1. Побудувати будь яку пряму через точку (0,0)
2. Побудувати звичайну функцію y = sin(x)
3. Побудувати пряму через першу і другу точку які була зчитана із файлу
4. Побудувати графік порогової функції активації
5. Побудувати пряму паралельну до осі Ox через точку яка має мінімальне значення координатиу y.
6. Побудувати пряму через точки (0,1), і точку (2, 7)
7. Побудувати будь яку пряму яка б розділяла дві групи точок
8. Побудувати пряму паралельну до осі Oy через точку яка має максимальне значення координати x
9. Побудувати дві будь які прямі різними кольорами так щоб їх було видно на графіку
10. Побудувати будь яку пряму так щоб всі точки знаходилися по один бік від неї