

深度學習

Lab Assingment 1

1. 目的

- (1)撰寫感知器學習演算法(Perceptron Learning Algorithm; PLA)
- (2)熟悉 Python 語法及相關函式庫(如：numpy，matplotlib 等)

2. 課程章節與實驗環境

2.1. 課程章節

Week 4: Linear Models 中 Perceptron Learning Algorithm

2.2. 實驗環境

限定使用 Python 程式語言

3. 實驗進行步驟及結果

3.1. 自行撰寫 PLA 程式，不可使用套裝軟體現成程式

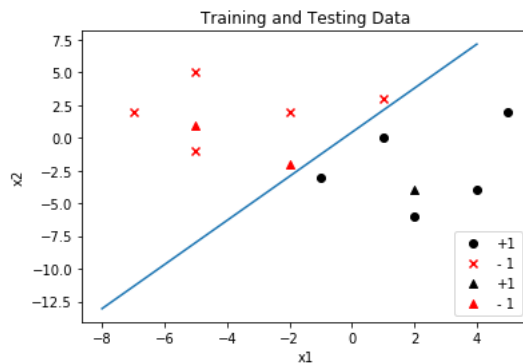
- 使用讀檔方式讀取資料
- 使用 Python 相關函式庫畫圖(此部分加分;最多額外加 20 分)

3.2. 訓練資料集/測試資料集

本實驗輸入資料為二維二分類訓練資料集，助教將提供具 200 筆之線性可分離(Linearly Separable)訓練資料集，其輸出類別值為 1 或 -1；另有 10 筆測試資料集(無輸出類別值)。

3.3. 程式結果要求

- 輸出分類線性函數 $w_0x_0 + w_1x_1 + w_2x_2 = 0$ 之 w_0, w_1, w_2 值
- 預測 10 筆測試資料之類別
- 畫圖(加分；最多 20 分)
使用 Python 相關函式庫(如：matplotlib)呈現二維平面上所有二分類訓練/測試資料及最後之線性函數： $w_0x_0 + w_1x_1 + w_2x_2 = 0$ ，如下示意圖(圖中資料點符號顏色，可自行選取)



(黑色•：

訓練資料類別值為 1；

紅色 x：

訓練資料類別值為-1；

黑色△/紅色△：

測試資料結果，其顏色需對應資料類別顏色，例如預測類別值為 1，以黑色△表示，預測類別值為-1，以紅色△表示。)

4. 說明

4.1. PLA 演算法

假設每筆訓練資料為 (\mathbf{x}, y) ，則 $\mathbf{x} = (x_0, x_1, x_2)$ 且 $x_0 \equiv 1$ ，而 $y = \pm 1$ 。

令 \hat{y}_k 為第 k 筆訓練資料估算值，則 $\hat{y}_k = \text{sign}(\mathbf{w} \bullet \mathbf{x}_k) = \pm 1$ 。

下表為 PLA 演算法：

Given a training data set:
 $D = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ // N training examples
Initialize all weights w_i to random values // $\mathbf{w} = (w_0, w_1, w_2)$;
UNTIL one of the termination conditions is met, DO
 FOR each training example $\mathbf{x}_k \in D$
 If $\text{sign}(\mathbf{w} \bullet \mathbf{x}_k) \neq y_k$ then $\mathbf{w} \leftarrow \mathbf{w} + y_k \mathbf{x}_k$

4.2. Net input n 之計算

Net input n 可用內積 (Inner Product) 或一般矩陣乘法計算：

– 內積

$$n = w_0x_0 + w_1x_1 + w_2x_2 = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \bullet \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \mathbf{w} \bullet \mathbf{x}$$

– 一般矩陣乘法

$$n = w_0x_0 + w_1x_1 + w_2x_2 = \begin{bmatrix} w_0 & w_1 & w_2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \mathbf{w}^T \mathbf{x}$$

4.3. PLA 停止條件

- 若是線性可分離 (Linearly Separable) 訓練資料，PLA 停止條件為所有訓練資料在某一世代 (Epoch) 均分類正確。
- 演算法最好還是給定最大世代 (Epoch) 數強迫停止，以免訓練資料為非線性可分離 (Linearly Inseparable) 時，致使 PLA 造成無窮迴圈，無法停止。

4.4. PLA 演算法變型，可具有學習率 (Learning Rate)：

$$\mathbf{w} \leftarrow \mathbf{w} + \eta y_k \mathbf{x}_k$$

學習率 η 之目的為控制 \mathbf{w} 更新調整幅度大小，其值一般為： $0 < \eta < 1$ 。