

지능형 시스템 HW01

Single-Layer Perceptron

정보컴퓨터공학과

201824602

최원준

1. 구현목적

숫자를 인식 할 수 있는 Single-Layer Perceptron을 설계하고 0~9까지의 숫자 데이터를 학습시킨 후 노이즈가 첨가된 숫자 데이터를 어느 정도의 정확성을 갖고 올바르게 판별하는지 알아본다. 그리고 암 환자와 관련된 데이터를 Single-Layer Perceptron을 제작하고 데이터를 학습시켜 주어진 2개의 임상자료가 어떤 암으로 판단되는지 알아보자.

2. 숫자 인식 Single-layer Perceptron

1) 숫자 데이터 디자인

Perceptron의 학습을 위하여 8x6크기(Width, Height)의 2차원 숫자 데이터를 Binary Image 형태로 제작한다.

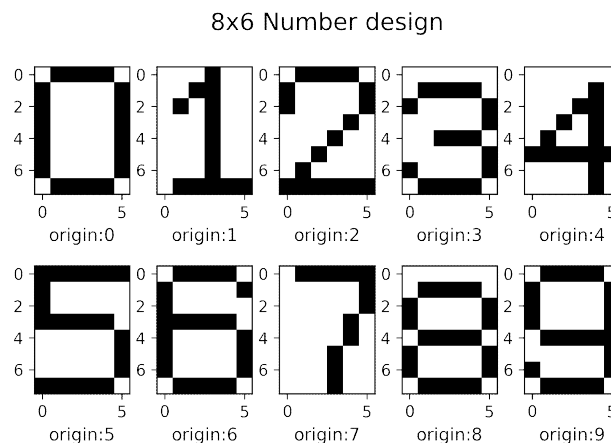


그림-1. Binary Image로 제작한 0~9 숫자 데이터

2) Desired Output 설정

임의의 0~9까지 숫자 데이터를 입력하여 판단되어 나오는 출력 값을 정해준다. 10개의 숫자 데이터를 구분 할 수 있도록 16가지의 경우를 구분 할 수 있는 4bits 2진수 형태로 Desired Output 값을 설정 해준다.

ex) 0 : (-1,-1,-1,-1), 1 : (-1,-1,-1, 1), ... , 8 : (1,-1,-1,-1), 9 : (1,-1,-1, 1)

3) 학습 단계

$$y(t) = f\left(\sum_i [x_i(t)w_i(t) - \theta]\right) \quad [1]$$

$$f(u) = \begin{cases} +1 & \text{if } u \geq 0, \\ -1 & \text{if } u < 0 \end{cases} \quad [2]$$

[1]의 식을 이용하여 현재 입력 $x(t)$ 와 현재 가중치 $w(t)$ 를 내적 하여 활성화함수 [2]에 넣어 설정해둔 Desired Output과 일치하는지 확인한다.

$$w_i(t+1) = w_i(t) + \eta[d(t) - y(t)]x_i(t) \quad 0 \leq i \leq N-1 \quad [3]$$

만약, 일치하지 않을 경우 해당 가중치 값을 [3]의 식을 사용하여 해당 가중치를 새롭게 갱신한다. 이 글에선 η (오차 반영률)을 1로 설정하여 빠르게 최적 가중치를 도출해 낼 수 있도록 제작하였다.

위 과정을 계속 반복하면서 각각의 입력 $x(t)$ 의 값이 미리 설정해둔 Desired Output과 모두 일치하면 가중치 갱신을 종료한다.

4) 인식 단계

기존 숫자 데이터에 노이즈를 무작위로 첨가하고 Perceptron이 제대로 인식되는지 확인하여 본다. 그림-2는 기존 숫자 데이터에 무작위 개수의 픽셀을 무작위 위치에 첨가하거나 제거하여 Perceptron에 인식시킨 결과이다. 초록색 라벨은 정확하게 숫자를 인식 하였으며 빨간색 라벨은 잘못된 값을 출력한 경우이다.

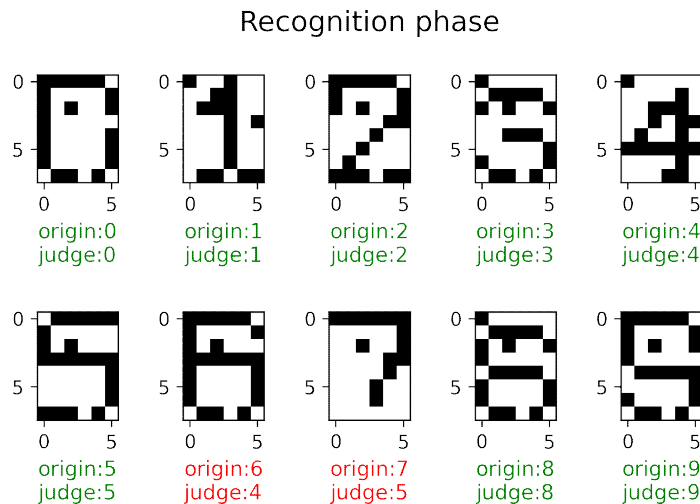


그림-2. 노이즈 숫자 데이터 인식
(초록:correct, 빨강:incorrect)

5) 정확도 측정

해당 Perceptron의 정확도를 측정하기 위해 기존 숫자 데이터에 무작위 노이즈가 첨가된 5개의 숫자를 각각 인식시켜 총 50개의 결과 중에 몇 개의 숫자가 제대로 인식되는지 확인해본다.

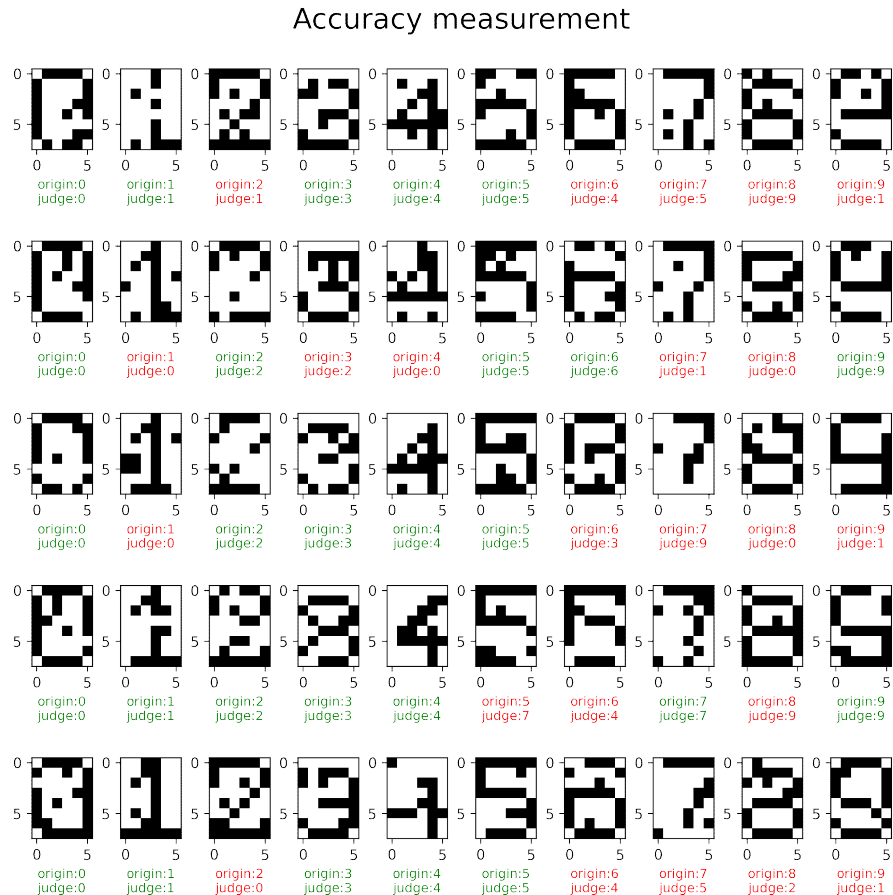


그림-3. 50개의 노이즈 숫자 데이터 학습 결과
(초록:correct, 빨강:incorrect)

그림-3은 무작위로 노이즈가 첨가된 숫자 데이터를 학습한 결과이며 50개의 데이터 중 27개의 데이터를 정확하게 인식하였다. 정확도는 54.0%를 보이고 있으며, 노이즈 위치에 따라 매번 학습 결과가 바뀌어 최대 약 70%까지의 정확도를 보이며 평균 약 65%의 정확도를 보이고 있다.

3. 암 판단 Single-layer Perceptron

암환자들로부터 얻은 임상결과 데이터 8개를 이전 숫자 인식 Perceptron과 마찬가지로 Desired Output를 설정해준다. 이때 Desired Output은 One-Hot Encoding 방식을 사용하여 구현 하였다. 하지만 학습 데이터가 너무 작아 학습을 시킬 때 마다 결과가 다르게 나왔으며, 그중 빈도가 제일 높은 진단 결과를 구하였다.

환자1(0.5, 0.7, 0.3, 0.5, 0.6)의 진단 결과는 간암

환자2(0.6, 0.7, 0.8, 0.9, 0.7)의 진단 결과는 폐암