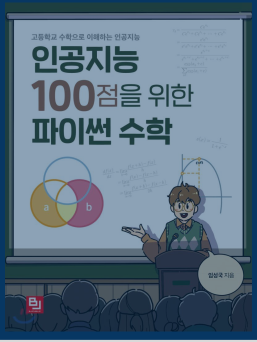
10. 손실함수

인공지능 100점을 위한 파이썬 수학



Contents

1. 손실함수란

01. 손실함수란

항등함수(identity function)

정확도는 학습을 진행할 때 큰 의미를 부여하지 못한다. 이유는 신경망의 결과와 정답의 차이를 보정해주지 못하기 때문입니다. 이를테면 신경망으로 3이 나온 결괏값이 0.51인 것과 0.99인 것의 차이가 없기 때문입니다. 그래서 정확도 대신 손실함수라는 것을 사용합니다. 조금 더 정확하게는 W1, W2, B1, B2의 값이 변화할 때 정확도는 연속적으로 변화하지 않고 단계적으로 변화하기 때문에 좋은 방향을 찾을 수 없다. 이것을 불연속성에 의한 미분불가라고 말합니다. 미분을 할 수 없고, 그래서 기울기를 찾을 수 없어서 학습하기 어렵기 때문에 최종 결과값이 아닌 확률값을 사용하고, 이것을 손실함수로 정의해서 사용합니다.

01. 손실함수란

항등함수(identity function)

y는 신경망을 통해서 나온 결괏값이고 t는 정답 레이블에 저장된 값입니다.

```
y = [0.01, 0.02, 0.05, 0.10, 0.12, 0.05, 0.05, 0.50, 0.00, 0.10]
t = [   0,    0,    0,    0,    0,    0,    0,    1,    0,    0]
```

01. 손실함수란

- 10 == 0.1 * 100 ?

0.1 을 100번 더하면 어떤값이 되는가?

2. 평균, 중간값, 표준편차, 분산

02. 평균, 중간값, 표준편차, 분산

평균, 기댓값

평균과 기댓값은 같은 내용을 다르게 표현하는 것.
표본을 모두 더한 후 그 갯수로 나눈것이 평균이고, 각 표본이 일어날 확률값에 표본의 값을 곱해서 전체에 대해 더한것이 기댓값이다.

$$m = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$E(X) = \sum_{i=1}^{N} x_i P(X = x_i)$$

02. 평균, 중간값, 표준편차, 분산

중간값 (median)

중간값은 데이터를 일렬로 늘어놓았을때 한 중앙에 위치하는 값이다. 데이터의 대표값으로 많이 사용되는 것이 평균과 중간값이다. 데이터에 이상값이 많은 경우 중간값이 자주 사용된다.

02. 평균, 중간값, 표준편차, 분산

표준편차, 분사

- 분산은 전체의 데이터가 평균과 얼마나 멀리 떨어진 곳에 분포하는지를 나타내는 값입니다. 모든 데이터의 값에서 평균을 뺀 다음 제곱한 값의 평균입니다.
- 표준편차는 분산에 제곱근을 씌워준 값입니다.

$$\sigma^2 = V(X) = E[(x-m)^2] = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2$$

$$\sigma = \sqrt{\sigma^2}$$

3. 평균제곱오차

03. 평균제곱오차

출력의 비중을 조정

평균제곱오차는 영어로 mean square error이며 줄여서 MSE로 사용합니다. 우리말로 풀어보면 ‘ 오차들의 제곱을 평균한 것 ’ 이라고 볼 수 있습니다.

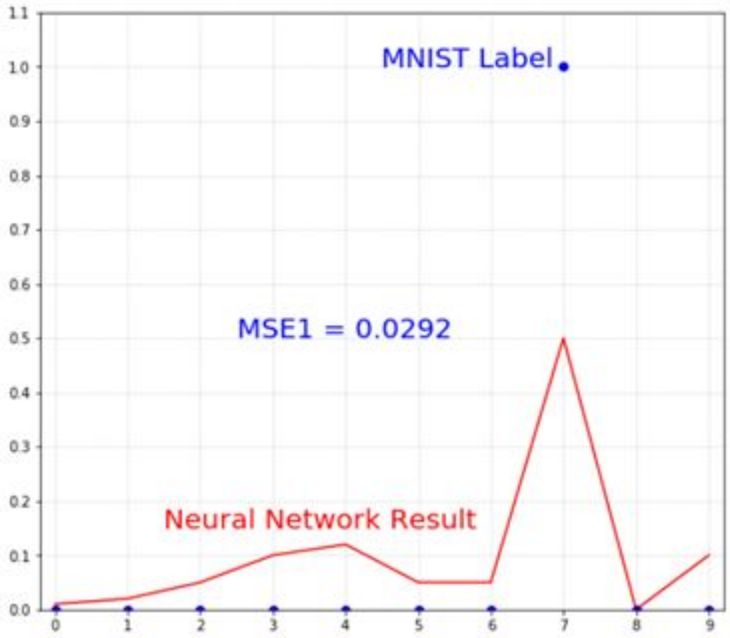
$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - t_i)^2$$

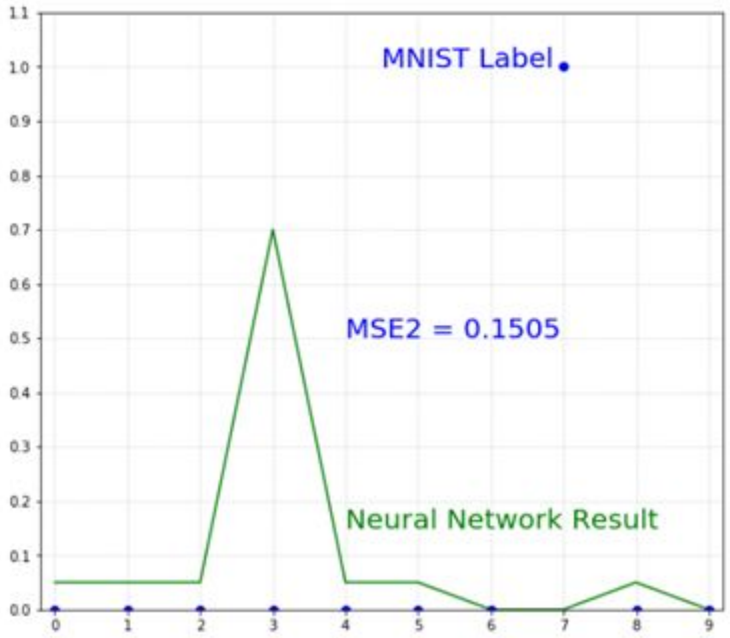
y_pred = [0.01 0.02 0.01 0.01 0.5 0 0 0 0 0]   - index : 4
t   = [0 0 0 0 1 0 0 0 0 0] - index : 4
MSE = (0.01**2 + 0.02**2 + 0.01**2 + 0.01**2 +(0.5-1)**2)/10

03. 평균제곱오차

출력의 비중을 조정

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - t_i)^2$$





4. 크로스엔트로피 오차

04. 크로스엔트로피 오차

출력의 비중을 조정

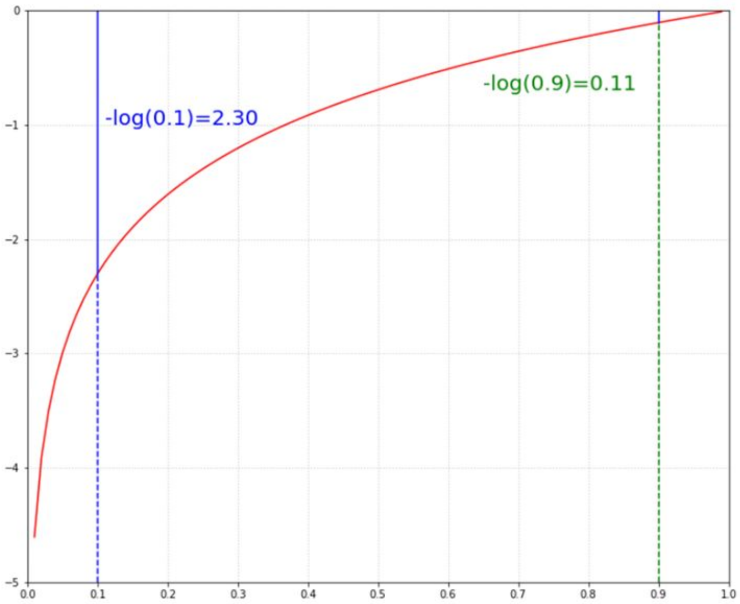
90%의 확률로 질 것이 뻔한 경기에 임했다고 가정을 해보겠습니다. I지역 조기축구회와 축구 명문대학인 A대 축구부가 경기를 했습니다. 여기서 A대 축구부가 이기면 아무도 놀라지 않습니다. 당연한 일이 발생했다고 여기고 그냥 넘어갈 것입니다. 반면 I지역 조기축구회가 이기면 어떻게 될까요? 아주 놀랄 만한 일이 발생한 것입니다. 이때의 놀라는 정도는 9:1에 비례하지 않습니다.

04. 크로스엔트로피 오차

출력의 비중을 조정

A대 축구부가 승리할 경우 : - log(0.9) = 0.1054
I지역 조기축구회가 승리할 경우 : - log(0.1) = 2.3026

A대 축구부가 승리할 때의 놀람의 정도는 0.1이고, I지역 조기축구회가 승리할 때의 놀람의 정도는 2.30이 됩니다. 이 수치가 곧 정보의 양이 됩니다. 조기축구회의 승리는 A대 축구부의 승리보다 21.85배(=2.3026/0.1054) 놀람을 가집니다. 여기서 살펴볼 것은 10%의 확률로 이기는 것이 10%의 정보량이나 놀람을 의미하지는 않는다는 것입니다. 더 많은 정보량을 가지게 됩니다.



04. 크로스엔트로피 오차

출력의 비중을 조정

Case I : 90% 유리 A 대 VS I 조기축구회	Case II : 50 대 50 승부 A 대 VS B 대
엔트로피 S = - { 0.9 × log(0.9) + 0.1 × log(0.1) } = 0.3251	엔트로피 S = - { 0.5 × log(0.5) + 0.5 × log(0.5) } = 0.6931

04. 크로스엔트로피 오차

MSE 와 CEE 비교