

Data Science(Machine Learning, Deep Learning)

- 실습 소스 및 참고

<심층 신경망 - DNN>

[DNN 데이터 준비 단계]

$$f(x) = wx + b$$

로지스틱 회귀

6:2:2의 비율로 학습 검증 테스트 데이터셋으로 나눴습니다

이전에는 7:3의 비율로 학습과 테스트 데이터셋으로 했음.(검증단계가 별도로 없었음)

[DNN 학습 코드 구현]

모델로 심층 신경망 만들어짐

활성함수(LeakyReLU 리키렐루)

모델 마지막에 시그모이드 함수 넣어준다. - 이진 분류

신경망 학습으로 Adam을 사용함.

BCE 함수를 손실 함수로 사용 ※ MSE Loss(MSE 손실 함수)와는 다름

BCE (Binary Cross-Entropy) 함수는 이진 분류(binary classification) 문제에서 주로 사용되는 손실 함수 중 하나

엔트로피(Entropy)는 정보 이론(Information Theory)에서 사용되는 중요한 개념 중 하나입니다. 엔트로피는 정보의 무질서도나 불확실성을 측정하는 지표로서, 주로 확률 분포의 불확실성을 표현하는 데 사용됩니다.

AUROC는 사이킷런에 있는 모듈, 이진 분류 모델의 성능을 평가하기 위해 사용 하는 모듈이다.

문제	타겟 값	마지막 계층의 활성함수	손실	예제
회귀	연속형 실수 벡터	없음	MSE 손실 함수	주어진 신상 정보를 바탕으로 연봉 예측하기
이진 분류	0 또는 1	시그모이드 함수	BCE 손실 함수	주어진 신상 정보를 바탕으로 성별 예측하기
분류	카테고리 또는 이산 값	소프트맥스 함수	CE 손실 함수	주어진 신상 정보를 바탕으로 직업군 예측하기

Cross-Entropy Loss(CE Loss) 또는 Cross-Entropy 손실 함수는 주로 분류 문제에서 사용되는 손실 함수 중 하나

[다중 클래스]

혼동 행렬

혼동 행렬(Confusion Matrix)은 분류 모델의 성능 평가를 위한 표 형태의 행렬

혼동 행렬은 다음 네 가지 항목으로 구성됩니다:

True Positive (TP): 실제 양성 클래스인 데이터를 양성으로 정확히 예측한 경우입니다. 즉, 모델이 양성이라고 예측하고 실제로 양성인 경우를 나타냅니다.

False Positive (FP): 실제 음성 클래스인 데이터를 양성으로 잘못 예측한 경우입니다. 즉, 모델이 양성이라고 예측했지만 실제로는 음성인 경우를 나타냅니다.

True Negative (TN): 실제 음성 클래스인 데이터를 음성으로 정확히 예측한 경우입니다. 모델이 음성이라고 예측하고 실제로 음성인 경우를 나타냅니다.

False Negative (FN): 실제 양성 클래스인 데이터를 음성으로 잘못 예측한 경우입니다. 모델이 음성이라고 예측했지만 실제로는 양성인 경우를 나타냅니다.

MNIST는 손으로 쓴 숫자(0부터 9까지) 이미지 데이터셋을 포함한 기계 학습 및 딥 러닝 연구를 위한 대표적인 데이터셋 중 하나

[Deep Classification(심층 분류)]

MNIST 데이터셋 분류 - Tensor

vaId - 경사하강법(미분) 필요 X

결과물로 Confusion_matrix로 인해서 혼동 행렬이 만들어짐