

PyTorch로 딥러닝 제대로 배우기

- 기초편 -

Part2. 머신러닝 개론

강사: 김 동 희

목차

▪ □ 머신 러닝 개념

- 1) Machine Learning
- 2) Prediction vs. Forecast
- 3) ML vs. DL
- 4) Task

▪ □ 머신 러닝 원리

- 1) 딥러닝 파이프 라인
- 2) 입력 데이터
- 3) 모델
- 4) 오차 측정 & 최적화

4. 오차 측정 & 최적화

❑ 목표 (Target)

- 데이터(X)를 통해 구하고자 하는 해답

❑ Supervised Learning

- 데이터(X)에 해당 Task 해결을 위한 정답 label이 존재하는 경우
- 이때, 오차 측정은 정답 label과 도출한 예측 값과 차이로 산출

❑ Unsupervised Learning

- 데이터(X)에 정답 label이 존재하지 않음
- 데이터간에 거리, 유사도 등을 비교하여 오차 측정을 수행

❑ Self-supervised Learning

- 데이터(X)에 정답 label이 존재하지 않지만, 강제로 정답 label을 만들어서 학습하는 방법
- 데이터(X)를 Rotation, Random Cut 등을 수행하여 만들어진 데이터(X')와 오차 측정을 하는 방법

4. 오차 측정 & 최적화

□ 손실 함수 (Loss Functions)

- L1Loss

$$l_n = |x_n - y_n|,$$

- MSELoss

$$l_n = (x_n - y_n)^2$$

- CrossEntropyLoss

$$l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^C \exp(x_{n,c})} \cdot 1\{y_n \neq \text{ignore_index}\}$$

- KLDivLoss

$$L(y_{\text{pred}}, y_{\text{true}}) = y_{\text{true}} \cdot \log \frac{y_{\text{true}}}{y_{\text{pred}}} = y_{\text{true}} \cdot (\log y_{\text{true}} - \log y_{\text{pred}})$$

- TripletMarginLoss

$$L(a, p, n) = \max\{d(a_i, p_i) - d(a_i, n_i) + \text{margin}, 0\}$$

4. 오차 측정 & 최적화

□ 최적화 함수 (Optimization Functions)

- Loss 함수로부터 산출된 오차값을 바탕으로 정답에 근사한 parameter를 찾아가게 해주는 함수
- Gradient를 구하고, 이를 특정 learning rate에 비례하여 parameter에 적용
- SGD
- Adam
- RMSprop

감사합니다.