



# DLPyTorch Ch.1

## 1.1 인공지능, 머신 러닝과 딥러닝

- 인공지능 > 머신러닝 > 딥러닝
- 머신러닝은 데이터를 인간이 전처리해서 컴퓨터에 제공해야하는 반면 딥러닝은 전처리 과정이 생략되고 컴퓨터가 스스로 진행함

## 1.2 머신 러닝

### 1. 머신러닝 학습 과정

- 데이터: 학습 모델에 사용되는 자료. ML에서는 훈련 데이터셋과 테스트 데이터셋 용도로 분리하여 사용함.
- 모델: 학습을 통해 얻은 일종의 '가설', 모델 선택 → 모델 학습과 평가 → 모델 업데이트의 반복.
- 

### 2. 머신러닝 학습 알고리즘

- 지도 학습(supervised learning): 정답 레이블을 알려주고 학습시키는 방법
  - 분류(classification)
  - 회귀(regression)
- 비지도 학습(unsupervised learning): 특성이 비슷한 데이터를 범주화하여 예측하는 방법
  - 군집(clustering)
  - 차원 축소(dimensionality reduction)

- 강화 학습(reinforcement learning): 행동에 대한 보상을 통해 특정 행동을 강화, 약화시키는 방법

## 1.3 딥러닝이란

- 인간의 신경망 원리를 기반으로 만들어진 머신 러닝 방법의 일종.
- 엄청난 수의 뉴런과 시냅스의 개념에서 착안. 복잡하게 연결된 뉴런들이 병렬 연결 되어 있음으로서 음성, 영상 인식 등의 처리를 가능하게 함.

## 1. 딥러닝 학습 과정

- 데이터 준비 → 모델 정의 → 모델 컴파일 (옵티마이저, 손실 함수 선택) → 모델 훈련 → 모델 예측
  - 데이터 준비: 제공된 데이터셋 사용, 공개 데이터셋 사용
  - 모델 정의: 신경망 생성 (은닉층 개수 설정)
  - 모델 컴파일: 활성화 함수, 손실 함수, 옵티마이저 선택. MSE, cross entropy 등 데이터 형태에 맞추어서 적절한 도구 선택
  - 모델 훈련: 배치와 에포크(epoch) 선택, 최적의 파라미터와 하이퍼파라미터 찾기.
  - 모델 예측: 모델에 데이터셋을 적용하여 실제로 예측을 진행
- 신경망 - 여러 은닉층을 가진 심층 신경망으로 이루어져 있음
- 역전파 - 오차의 편미분을 통해 가중치 값을 업데이트하여 모델의 성능을 높임.

## 2. 딥러닝 학습 알고리즘

- 지도 학습
  - 이미지 분류
    - CNN: 합성곱 신경망. 이미지 분류, 인식, 분할로 분류.
  - 시계열 데이터 분석
    - RNN: 순환 신경망, 시간에 따른 데이터 분석할 때 이용

- LSTM: RNN에서 기울기 소멸 문제를 해결하여 게이트를 추가한 신경망.
- 비지도 학습
  - 워드 임베딩: 자연어를 변환하여 벡터로 표현.
  - 군집: GMM, SOM
  - 차원 축소: AutoEncoder, PCA
- 전이 학습: 사전에 학습이 완료된 모델을 가지고 우리가 원하는 학습에 미세 조정 기법을 이용해 학습시키는 방법. BERT 등
- 강화 학습: MDP