

Chapter1. 한 눈에 보는 머신러닝

1. 지도학습: 훈련 데이터와 그에 따른 정답을 함께 제공하는 학습 방법
 - 분류(Classification)
 - 회귀(Regression)
2. 비지도학습: 훈련 데이터에 레이블이 없어 시스템이 아무런 도움 없이 학습해야 하는 방법
 - 군집(Clustering) | 데이터들을 그룹화(ex. 블로그 방문자 사이의 연결고리를 찾아 그룹화)
 - 시각화 | 레이블이 없는 대규모 고차원 데이터를 도식화가 가능한 2D나 3D로 표현
 - 차원 축소 | 상관관계가 있는 여러 특성을 하나로 합치는 작업(Feature Extraction)
 - 이상치 탐지 | 데이터셋에서 이상한 값을 자동으로 제거하고, 새로운 샘플이 들어왔을 때 정상 데이터인지 이상치인지 판단
 - 특이치 탐지 | 훈련 세트에 있는 모든 샘플과 달라 보이는 새로운 샘플 탐지 → 아주 깨끗한 훈련 세트가 필요
 - 연관 규칙 학습 | 대량의 데이터에서 특성 간의 흥미로운 관계를 찾는 것(ex. 슈퍼에서 바비큐 소스와 감자를 구매한 사람이 스테이크도 구매하는 경향이 있음을 발견)
3. 준지도 학습: 일부만 레이블이 있는 데이터를 다루는 학습 방법
4. 강화 학습: 에이전트가 환경을 관찰해서 행동을 실행함으로써 리워드와 패널티를 받고, 가장 큰 보상을 얻기 위해 정책(Policy) 전략을 스스로 학습하는 방법

Chapter4. 모델 훈련

1. 선형 회귀(Linear Regression): 데이터를 일반화하는 일차함수의 기울기와 y절편을 찾아나가는 과정
 - 정규방정식(행렬)의 해를 찾는다.
 - 특이값분해(Singular Value Decomposition)를 이용해 파라미터를 찾는다.
 - 경사하강법(Gradient Descent): 비용 함수의 최솟값을 찾아나가는 과정 | 경사하강법을 사용할 때엔 StandardScaler를 사용하자.
 - 배치 경사하강법: 비용함수의 그레디언트 벡터를 매 스텝에서 전체 훈련세트 X에 대해 계산
- $$\theta' = \theta - \eta \nabla_{\theta} MSE(\theta)$$
- 확률적 경사하강법: 매 스텝에서 한 개의 샘플을 무작위로 선택하고 그 하나의 샘플에 대한 그레디언트를 계산
 - 배치 경사하강법에 비해 속도가 빠름
 - 미니배치 경사하강법: 미니배치라는 작은 샘플 세트에 대해 그레디언트를 계산
2. 다항 회귀(Polynomial Regression): 데이터를 일반화하는 n차 함수의 파라미터를 찾아나가는 과정
3. 규제가 있는 선형 모델
 - 릿지 회귀(티호노프 규제): 비용함수에 규제항을 추가한 것. α 가 0이면 선형회귀와 같아지고, α 가 매우 크면 평균을 지나는 수평선이 된다.

$$J(\theta) = MSE(\theta) + \alpha \frac{1}{2} \sum_{i=1}^n \theta_i^2$$

- 라쏘 회귀: 릿지 회귀와 비슷하나 규제항이 θ 의 제곱이 아닌 절댓값으로 표현됨.

$$J(\theta) = MSE(\theta) + \alpha \frac{1}{2} \sum_{i=1}^n |\theta_i|$$

- 엘라스틱넷(Elastic Net): 릿지 회귀와 라쏘 회귀를 절충한 모델. r 이 0이면 릿지 회귀와 같고, r 이 1이면 라쏘 회귀와 같다.

$$J(\theta) = MSE(\theta) + r\alpha \sum_{i=1}^n |\theta_i| + \frac{1-r}{2}\alpha \sum_{i=1}^n \theta_i^2$$

4. 로지스틱 회귀(Logistic Regression): 시그모이드 함수를 통해 확률을 추정하는 회귀 방법
5. 소프트맥스 회귀(Softmax Regression): 여러 개의 이진분류기(로지스틱 회귀 등)를 훈련시켜 연결하지 않고 직접 다중 클래스를 지원하도록 일반화될 수 있도록 한 것. 다항 로지스틱 회귀라고도 함.