

| | |
|-------|---|
| 교육 제목 | 데이터 기반 인공지능 시스템 엔지니어 양성 과정_ 머신러닝 |
| 교육 일시 | 2021년 10월 13일 |
| 교육 장소 | YGL C-6 학과장 & 자택(디스코드 이용한 온라인) |
| 교육 내용 | |
| | <p>지난 시간 Review & 미분계속</p> <p>1. 도함수</p> <p>$f(x)$와 $g(x)$가 미분가능한 함수이고 k가 상수일 때 다음이 성립한다.</p> $(1) (kf)'(x) = kf'(x)$ $(2) (f+g)'(x) = f'(x) + g'(x)$ $(3) (f-g)'(x) = f'(x) - g'(x)$ $(4) (fg)'(x) = f'(x)g(x) + f(x)g'(x)$ $(5) \left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$ <p>n° 임의의 정수일 때 다음이 성립한다.</p> $\frac{d}{dx}(x^n) = nx^{n-1}$ <p>2. 연쇄법칙(합성함수의 미분법)</p> <p>함수 $g(x)$가 미분가능하고 $f(x)$가 $g(x)$의 치역을 포함하는 영역에서 미분가능하면 합성함수 $y = (f \circ g)(x)$도 미분가능하고, $y = (f \circ g)(x) = f(g(x))$에서 $u = g(x)$라 놓으면</p> $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$ <p>또는</p> $(f \circ g)'(x) = f'(g(x))g'(x)$ <p>3. 음함수(implicit function)의 미분</p> <p>음함수 $F(x, y) = 0$에서 $\frac{dy}{dx}$를 구할 때는 y를 x의 함수로 보고 양변을 x에 관하여 미분한 다음 $\frac{dy}{dx}$를 좌변으로 분리하면 된다.</p> <p>4. 매개변수함수의 미분</p> <p>$x = f(t)$와 $y = g(t)$가 t에 관하여 미분가능이고 $f'(t) \neq 0^\circ$면,</p> $\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{g'(t)}{f'(t)}$ <p>5. 삼각함수의 미분</p> $(1) (\sin x)' = \cos x \quad (2) (\cos x)' = -\sin x$ <p>6. 로그함수의 미분</p> |

$$(1) (\log_a x)' = \frac{1}{x} \log_a e$$

$$(2) (\ln x)' = \frac{1}{x}$$

7. 지수함수의 미분

$$(1) (a^x)' = a^x \ln a$$

$$(2) (e^x)' = e^x$$

선형회귀(Linear regression)

1. 단순회귀(Simple linear regression)

$$f(x_i) = Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \text{ for } i = 1, 2, \dots, n$$

- $\epsilon_i \sim^{i.i.d} N(0, \sigma^2), \forall i$
- Independence, Normality, Homoscedasticity

2. 최소제곱추정량(OLS : Ordinary least square estimation)

$$\text{Linear model, } \hat{y}(x_i) = \hat{\theta}_0 + \hat{\theta}_1 x_i$$

3. 손실함수(Loss function, J)

$$J(\theta) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \hat{\theta}_0 - \hat{\theta}_1 x_i)^2 = \sum_{i=1}^n e_i^2$$

- $\hat{y}_i = \hat{\theta}_0 + \hat{\theta}_1 x_i$

- Sum of Squares of the Errors (SSE) = $\sum_i^n e_i^2$

- Goal is to solve for $\hat{\theta}_0$ and $\hat{\theta}_1$ to minimize the objective function.

$$\hat{\theta}_0, \hat{\theta}_1 = \operatorname{argmin}_{\theta_0, \theta_1} \sum_i^n e_i^2 = \operatorname{argmin}_{\theta_0, \theta_1} J(\theta)$$

$$\cdot \frac{\partial J(\theta)}{\partial \hat{\theta}_0} = 0, \quad \hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{x}$$

$$\cdot \frac{\partial J(\theta)}{\partial \hat{\theta}_1} = 0, \quad \hat{\theta}_1 = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_i^n (x_i - \bar{x})^2}$$

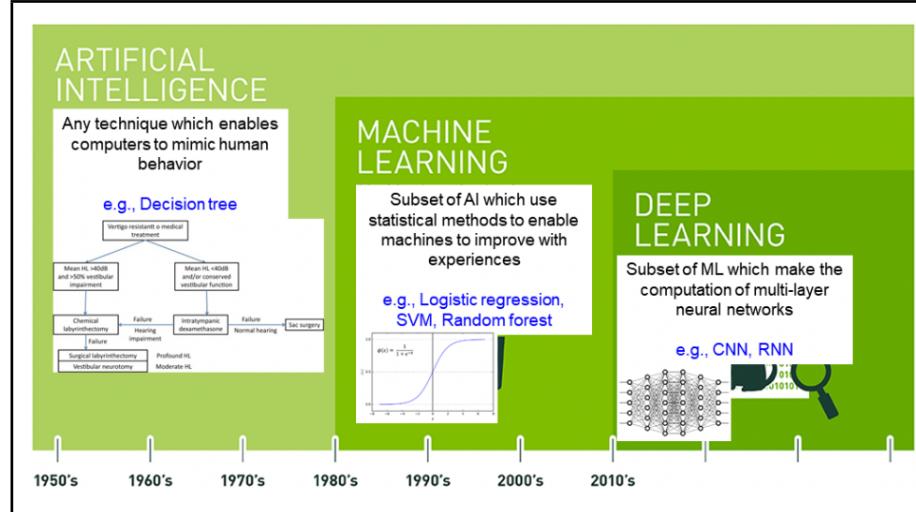
$$\cdot \sum_i^n (x_i - \bar{x})(y_i - \bar{y}) = S_{xy}$$

$$\cdot \sum_i^n (x_i - \bar{x})^2 = S_{xx}$$

ML Modelling process

머신러닝이란 일련의 수학적 과정을 통해 Error를 최소화 할 수 있는 Omega_0 와 Omega_1... 을 찾는 것이다. (기계는 패턴을 인식한다.)

1. AI vs ML vs DL



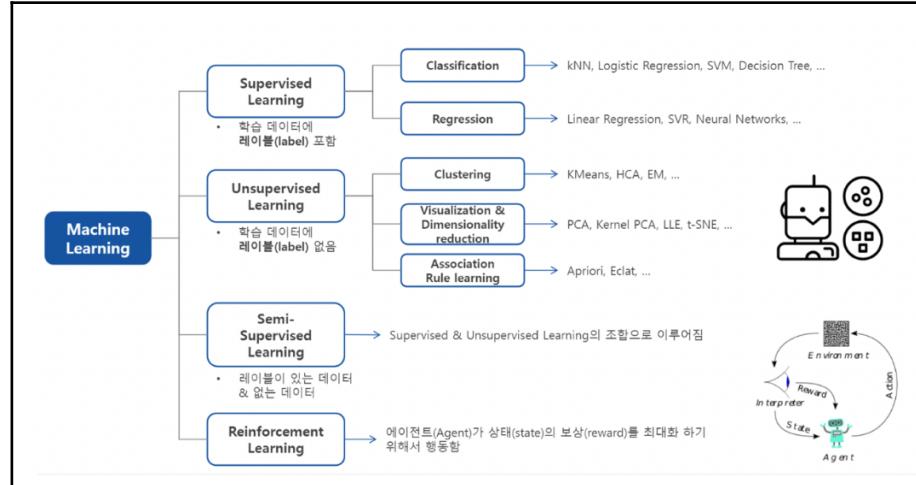
오후

Image classification 능력은 2015년 이후 인간의 능력을 앞서기 시작했다.

2016년 3월 알파고는 이세돌을 바둑에서 이겼다.

2017년 흉부 X-ray 분석 능력도 전문가를 앞서기 시작했다.

2. 기계학습의 분류



레이블 = 정답 : 데이터에 정답이 있고 이를 예측하는 경우 지도 학습, 데이터에 정답이 없고 패턴을 도출하는 경우 비지도 학습

모델링 과정(ML Modelling process)

1. Training set and test set

- ML 모델의 성능평가를 위해서 자료를 분할
- **Training set:** 모델의 알고리즘 learning, 모델에 사용될 feature들을 결정, 초매개변수 조절 (약 전체 자료수의 70%로 설정)
 - **Training set:** 모델의 알고리즘 learning
 - **Validation set:** 모델에 사용될 feature들을 결정, 초매개변수 조절, 과적합(Overfitting) 방지
- **Test set:** 최종 선택된 모델의 성능평가 (약 전체 자료수의 30%로 설정), 자료의 수가 적을 경우 생략 가능

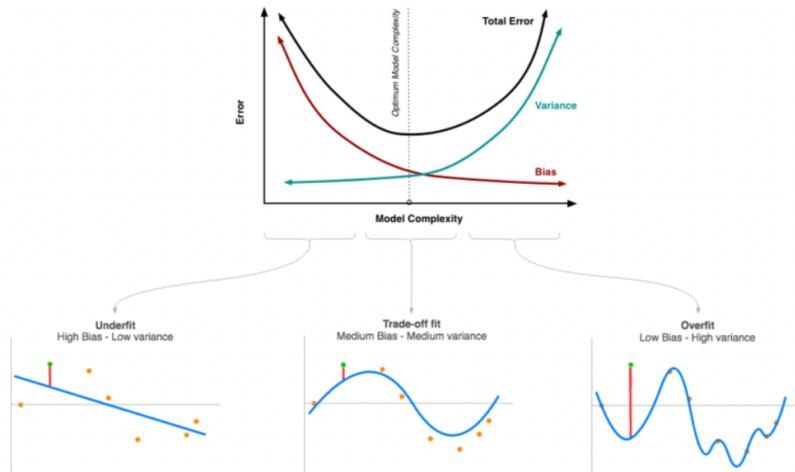
2. ML 모델의 치우침(Bias)과 분산(variance)

“ 예측값들과 정답이 대체로 멀리 떨어져 있으면 결과의 편향(bias)이 높다고 말하고, 예측값들이 자기들끼리 대체로 멀리 흩어져 있으면 결과의 분산(variance)이 높다고 말합니다.

과소적합(underfitting)은 치우침이 높은 모델이 발생하기 쉬움

과적합(overfitting)은 분산이 높은 모델이 발생하기 쉬움

Total Error가 최소화 되는 지점을 찾아 초매개변수를 정해야 함



3. 기계학습 모델 평가 : k-fold 교차검증(k-fold cross validation(CV))

- k-fold 교차 검증 (일명 k-fold CV)은 훈련 데이터를 동일한 크기의 k 그룹 (k-fold)으로 무작위로 나누는 리샘플링 방법
- k-fold CV 추정치는 k 테스트 오류를 평균화하여 계산.

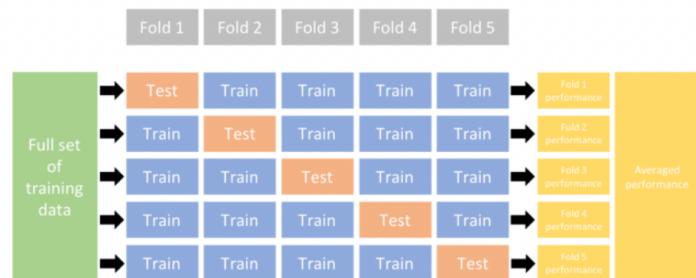


Figure 2.4: Illustration of the k-fold cross validation process.