

A.I 학생 포트폴리오



동양미래대학교 20191768 -신윤규-

목차

인공지능& 머신러닝	1
딥러닝	4
인공신경망(ANN)	6
입출력	9
신경망연산	12
회귀와 분류	12
오차역전파.....	18
이진이항분류.....	19

인공지능이란?

인공지능(Artificial Intelligence) 인간의 두뇌 기능을 기계로 실현하는 것을 목적으로 한다.

약인공지능

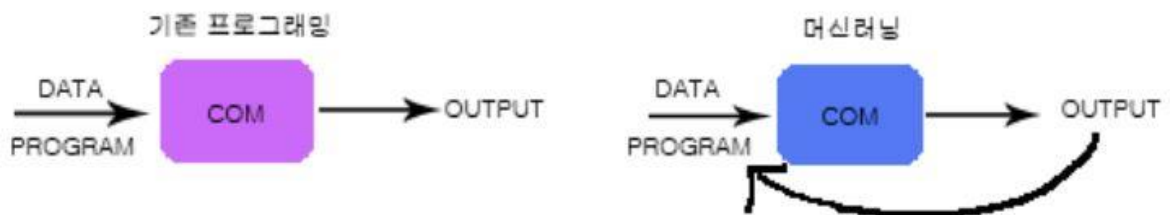
유용한 도구로 사용하기 위해 설계된 인공지능

강 인공지능

인간을 모방한 인공지능

머신러닝?

머신러닝 = 기계학습 = 기계가 자기 스스로 학습하는 것



머신러닝의 종류

지도학습(Supervised learning) = 정답있는 예측

올바른 입력& 출력의 쌍으로 구성된 정답의 훈련데이터(labeled data)로 IN/OUTPUT 간 함수의 학습이다.

K 최근접이웃(K-Nearest neighbor)

선형회귀(Linear regression)

로지스틱회귀(Logestic regression)

서포트벡터머신(support vectormachin) = svm

결정트리(decision Tree) & 랜덤 포레스트(Random Forest)

비지도(자율) 학습(Unsupervised learning)

정답이 없을 때 훈련데이터(unlabeled data)를 사용해 데이터내 숨어있는 관계를 찾는 방법이다.

군집(clustering)

K 평균(k-means)

계층군집분석(Hierarchical cluster analysis) = HCA

기댓값 최대화(expectation maximization)

시각화(visualization)&자원축소(Dimensionality reduction)

주성분분석(Principal component analysis) = PCA

커널(kernel PCA)

지역적 선형 임베딩(Logically – linear Embedding)

연관규칙학습(Association rule learning)

어프라이어리(Apriori)

이클랫(Eclat)

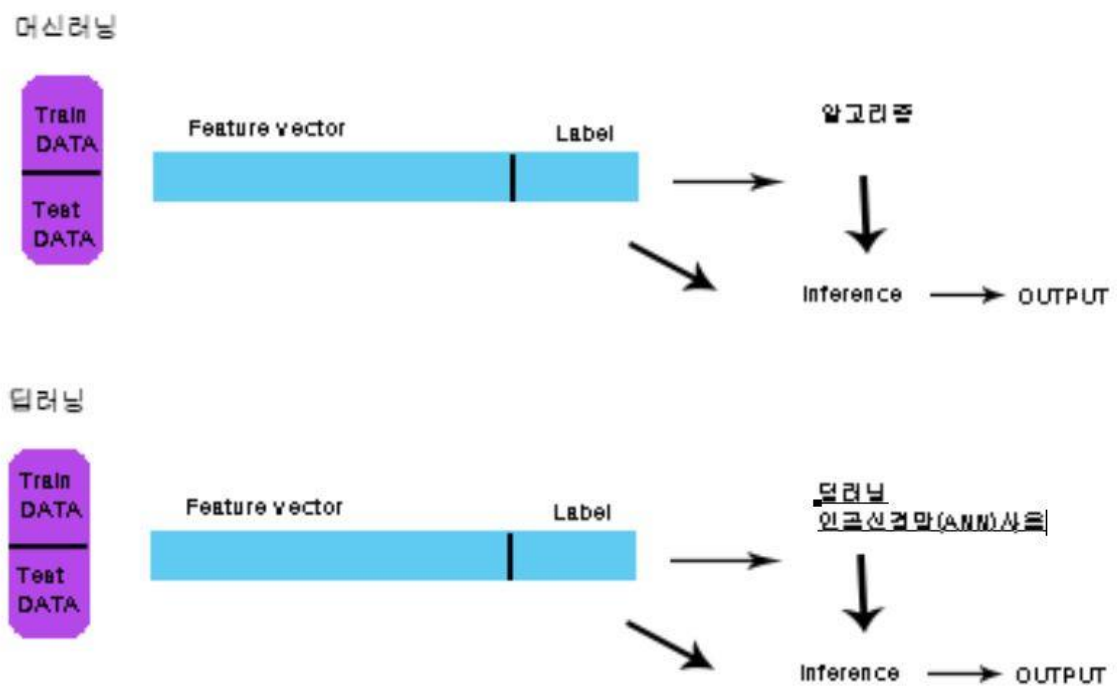
강화학습(reinforcement learning)

Ex) 딥마이닝 알파고 / 자동 GAME 분야

딥 러닝?

층을 만들어 보다 정확한 규칙을 만들어 나가는 것이다.

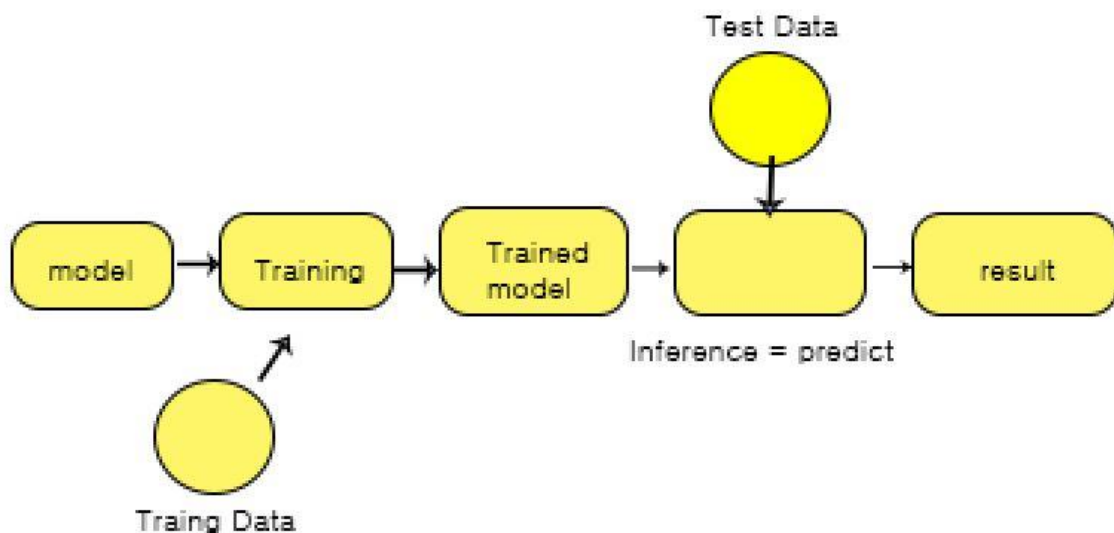
머신러닝 & 딥러닝



머신러닝 & 딥러닝 차이

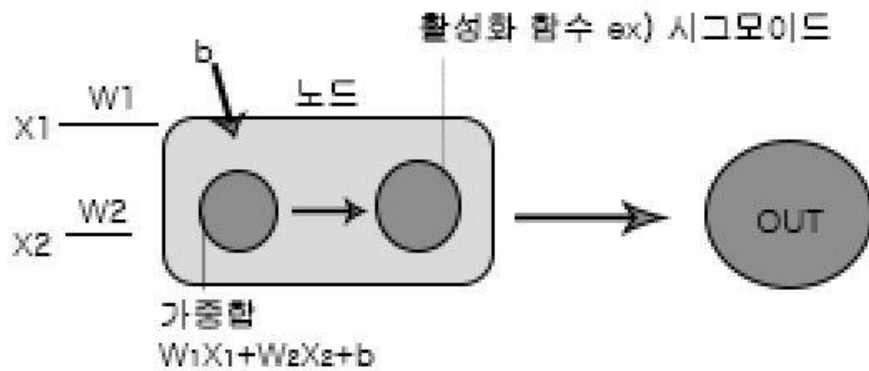
	기계학습	딥러닝
데이터 의존	중소형 데이터 세트	큰 데이터 세트
하드웨어 의존	감소형 머신	GPU 필요상당한 양의 행렬곱셈 실행
기능공학	데이터 나타내는 기능이해	기능을 이해하지 못한다.
실행시간	몇 분에서 몇 시간	최대 몇 주

딥러닝 과정



인공신경망(ANN)

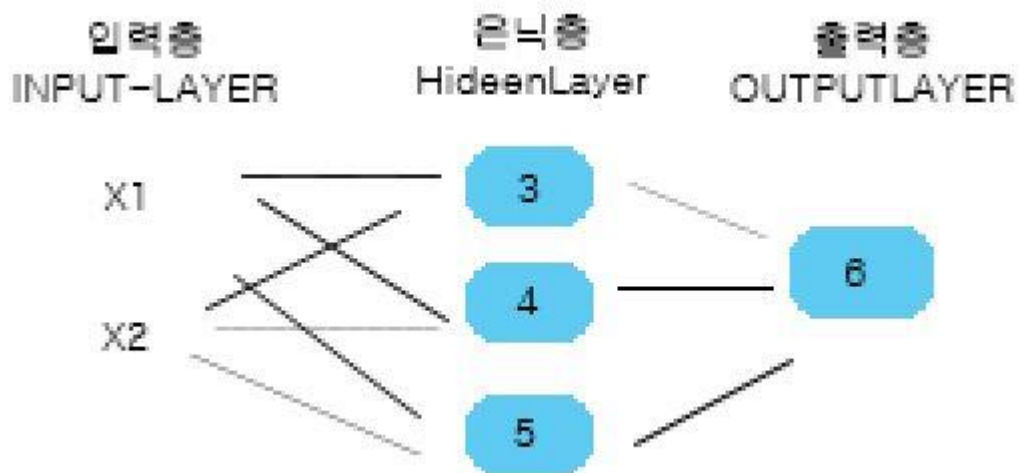
인공신경망은 시냅스의 결합으로 노드의 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결을 할 수 있는 모델들을 일컫는다..



MLP

입력층 & 출력층 - 다수신호 입력받아 하나신호 출력

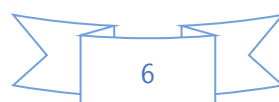
중간의 은닉층 - 여러 개 층 연결하여 하나의 신경망 구성



DNN = 심층 신경망

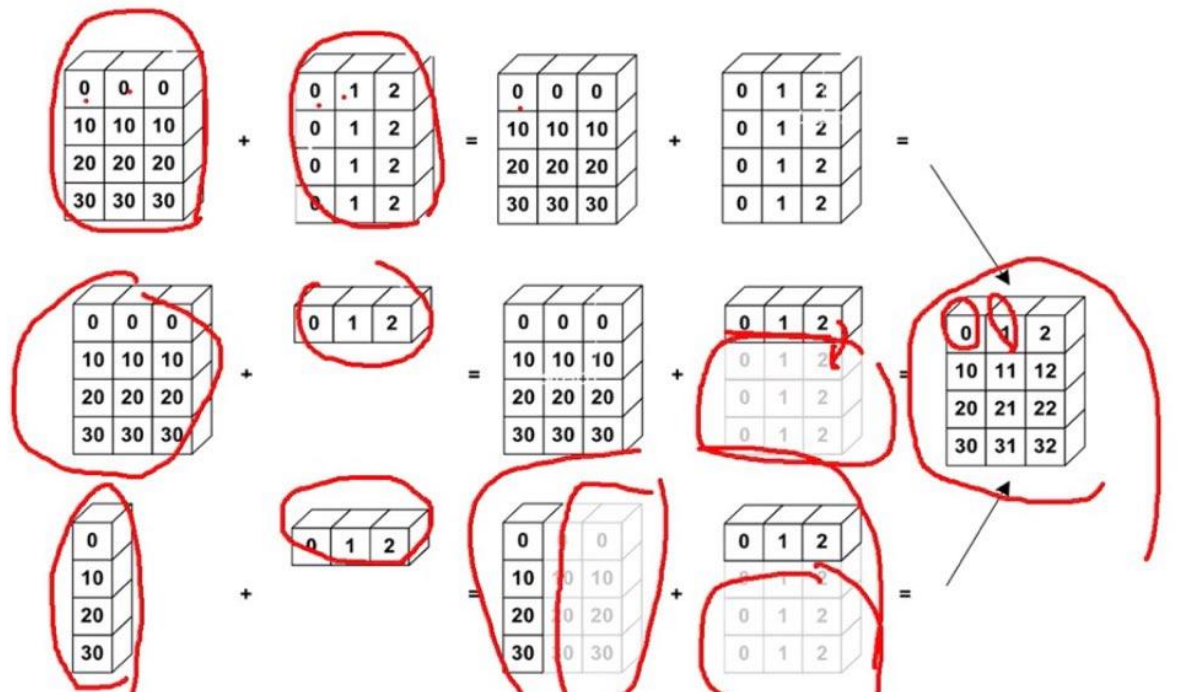
MLP 에 은닉층 2 개 이상 사용

OutPut layer 다수



브로딩 캐스팅 코드

이름 쓰지 않은 변수의 shape를 맞추



텐서플로우



다양한 작업에 대해 데이터 흐름 프로그래밍을 위한 오픈소스 소프트웨어 라이브러리이며 2015 년 11 월 9 일 공개되었다.

Colab



Colaboratory 라 부르고 줄여서 colab 이라 부르며 브라우저에서 Python 을 작성 및 실행할 수 있다. Colab 은 장점은 세가지로 구분을 할 수가 있다.

- 구성이 필요하지 않음
- GPU 무료 액세스
- 간편한 공유

케라스 딥러닝 구현



Define – 딥러닝 모델 만드는 것

Compile – 주요 훈련방법 설정

- 최적화 방법(Optimizers)
- 손실 함수(losses)
- 훈련 모니터링 지표(metrics)

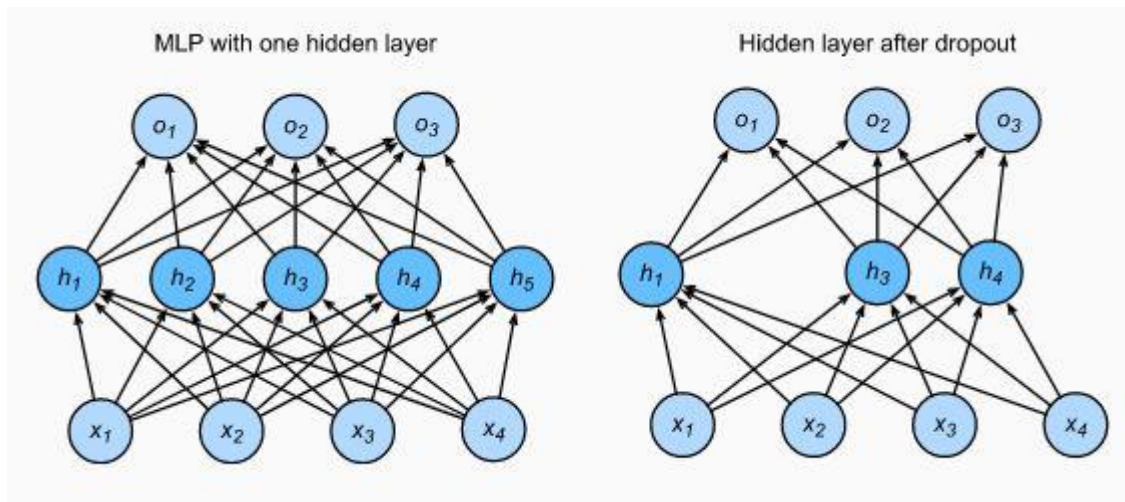
Fit – 훈련

Evaluate – 텍스트 데이터 평가

Predict – 정답 예측

드롭아웃

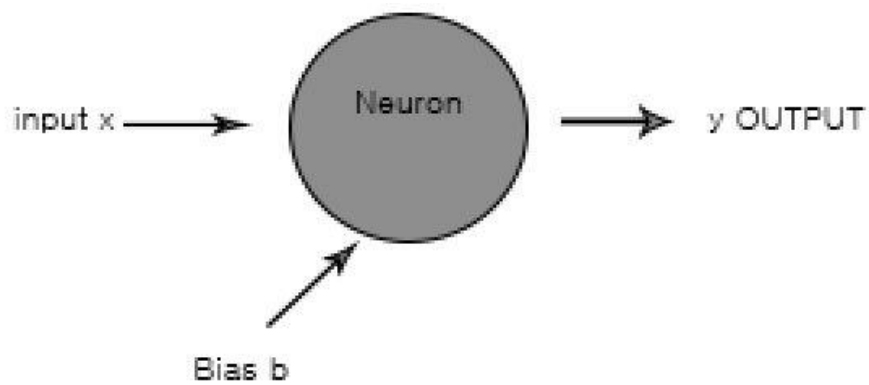
2012 년 개발 층에 결과 값을 일정 비율로 제거하는 방법이다.



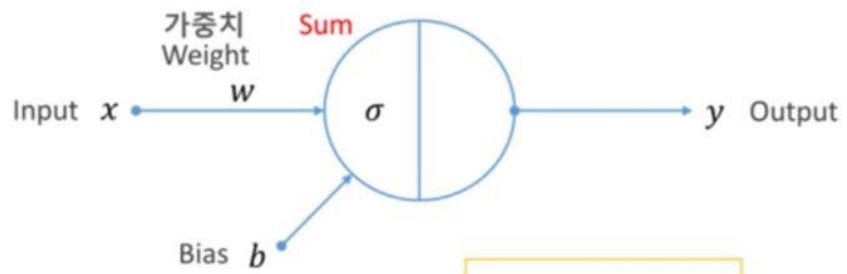
입출력

편향(bias)

편향을 조정해 출력을 맞춘다.

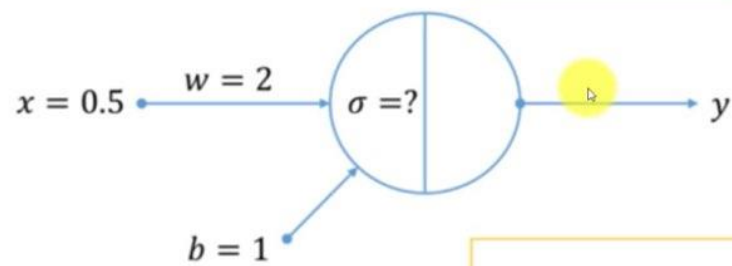


- 뉴런 식



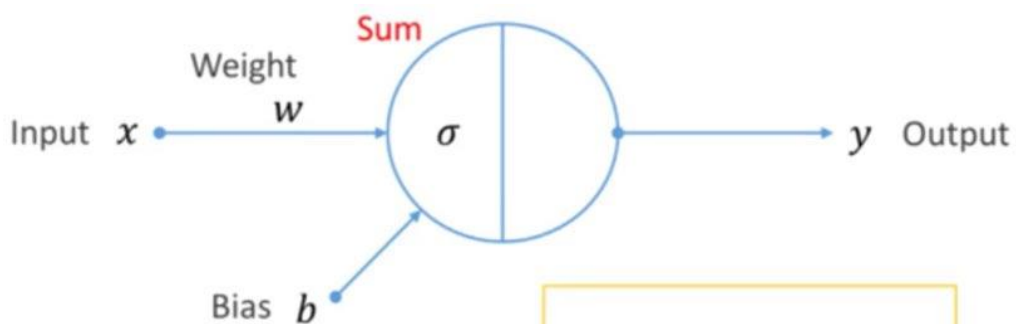
$$\sigma = w \cdot x + b$$

- 가중치와 편향



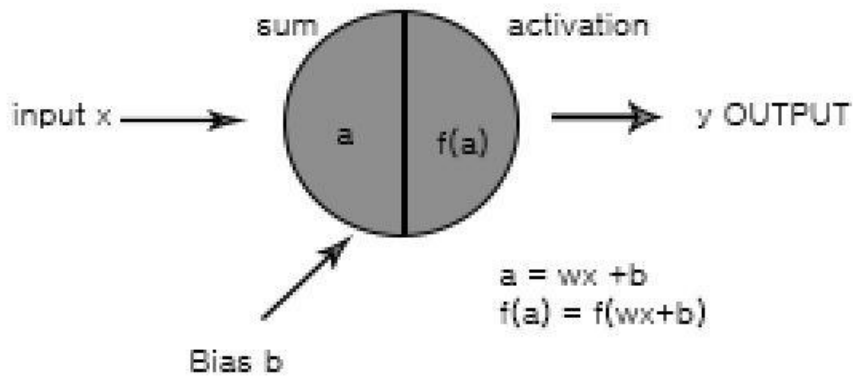
$$\begin{aligned}\sigma &= w \cdot x + b \\ &= 2 \cdot 0.5 + 1 \\ &= 2\end{aligned}$$

행렬곱 연산



$$\begin{aligned}\sigma &= w \cdot x + b \\ &= [w \ b] \begin{bmatrix} x \\ 1 \end{bmatrix}\end{aligned}$$

활성화 함수



RELU(Rectified Linear Unit)

0 이하는 모두 0 으로 한 함수

$\text{Max}(x, 0)$

Sigmoid

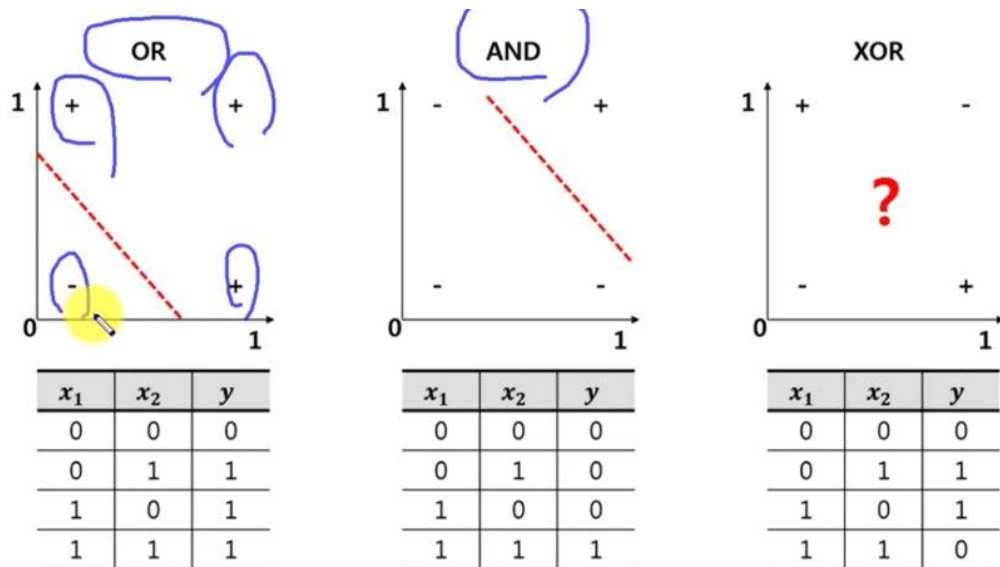
S 자 형태의 곡선

활성화 함수 & 편향

결과 값이 임계 값 역할

신경망연산

XOR 문제



회귀와 분류

회귀 모델

연속적 값 예측 EX) 캘리포니아 주택가격 얼마?

사용자가 이광고를 클릭할 확률 얼마

분류 모델

불연속적 값 예측 ex) 주어진 이메일 스팸? Or 스팸 no

이미지가 고양이 or 햄스터?

회귀 분석

연속형 변수에 두 변수 사이에 모형을 구한 뒤 적합도 측정

단순선형회귀 분석($H(x) = wx + b$)

입력: 특징이 하나/출력: 하나의 값

Ex) 키 몸무게 측정

다중선형회귀분석($y = w_1x_1 + \dots + w_nx_n + b$)

입력: 특징이 여러 개/출력: 하나의 값

Ex) 역세권&아파트 평수, 주소로 아파트 값 추정

로지스틱 회귀(이진분류) - $H(x) = Wx + b$ // w =가중치, b =편향

입력: 하나 or 여러 개 / 출력 0 or 1

Ex) 타이타닉 승객정보로 죽음 추정

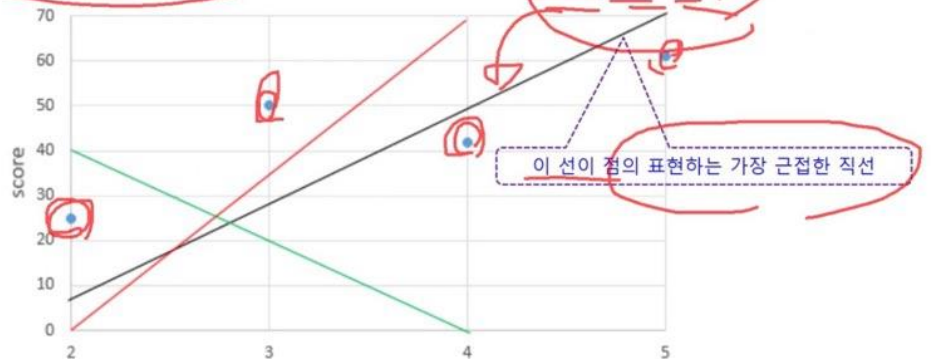
가설(Hypothesis)

가중치&편향 / 기울기&절편

- $H(x)$ 에서 H 는 Hypothesis를 의미

$$H(x) = Wx + b$$

W : 기울기, 가중치
 b : 절편, 편향



손실함수(Loss function)

MSE(Mean Square Error 평균 제곱오차)

Categorical crossentropy

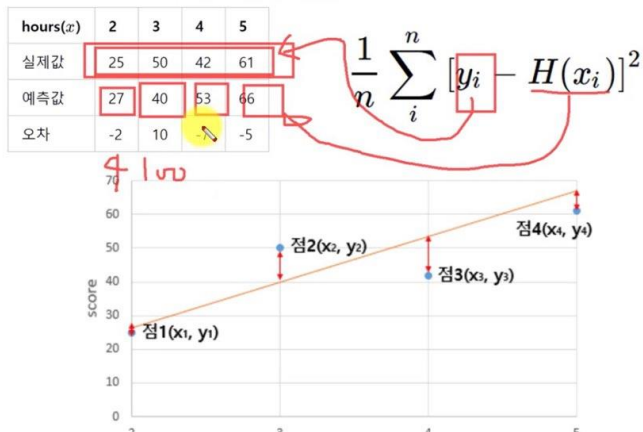
Sparse Categorical crossentropy

- 평균 제곱 오차(Mean Squared Error, MSE) 등을 사용

$$\frac{1}{n} \sum_i [y_i - H(x_i)]^2$$

실제 값 예측 값

- W : 13 b : 1로 예측한다면 $y = 13x + 1$ 직선이 예측한 함수로 예측 값을 추정



- 평균 제곱 오차를 W 와 b 에 의한 비용 함수(Cost function)로 재정의

$$\text{cost}(W, b) = \frac{1}{n} \sum_i^n [y_i - H(x_i)]^2$$

- 모든 점들과의 오차가 클수록 평균 제곱 오차는 커지며,
 - 오차가 작아질수록 평균 제곱 오차는 작아짐

- 평균 제곱 오차

- $\text{cost}(W, b)$ 를 최소가 되게 만드는 W 와 b 를 구하면
 - 결과적으로 y 와 x 의 관계를 가장 잘 나타내는 직선을 그릴 수 있게 됨

$$W, b \rightarrow \text{minimize cost}(W, b)$$

옵티마이저 : 최적화 과정

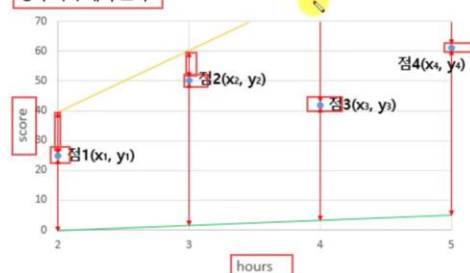
- 최적화 알고리즘
- 적절한 w, b 찾는 과정

경사하강법(Gradient Descent)

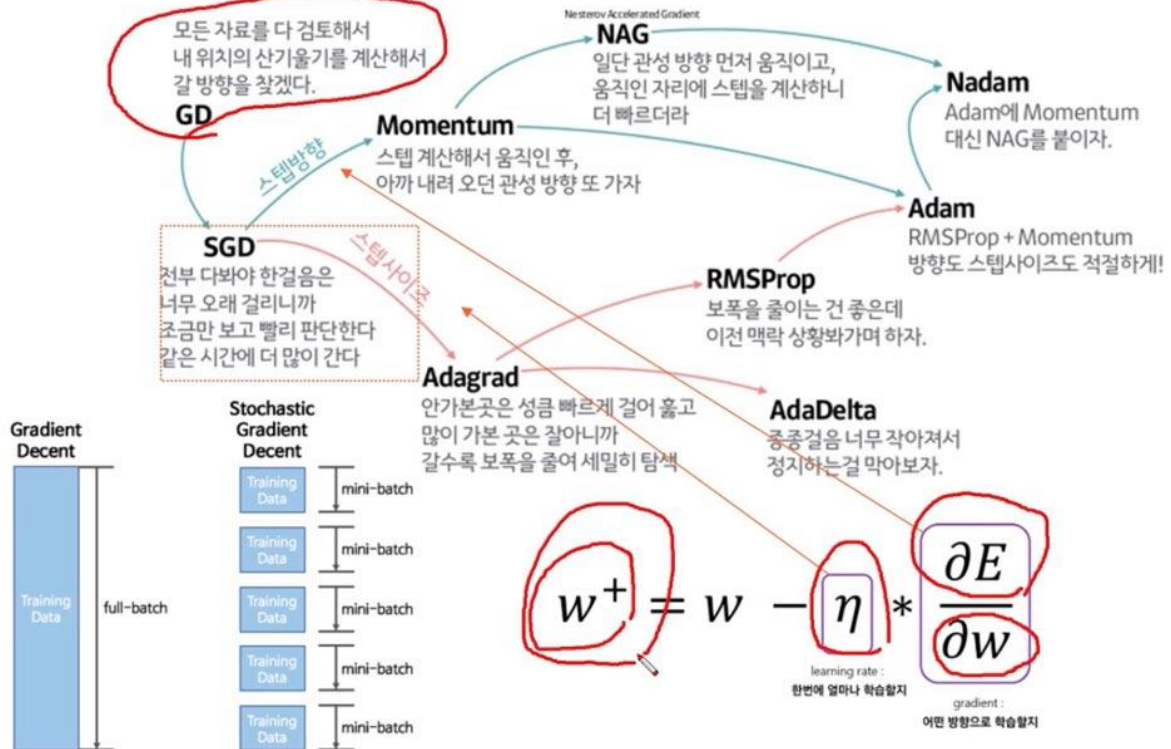
- 비용함수 값을 최소화하는 $w \& b$ 찾기
- 경사 따라 내려오기

- 머신 러닝에서 학습(training)

- 최적화 알고리즘(Optimizer algorithms)
- 적절한 W 와 b 를 찾아내는 과정
 - Gradient Descent(경사 하강법)
 - 비용 함수(Cost Function)의 값을 최소로 하는 W 와 b 를 찾는 방법
 - 경사 따라 내려 오기

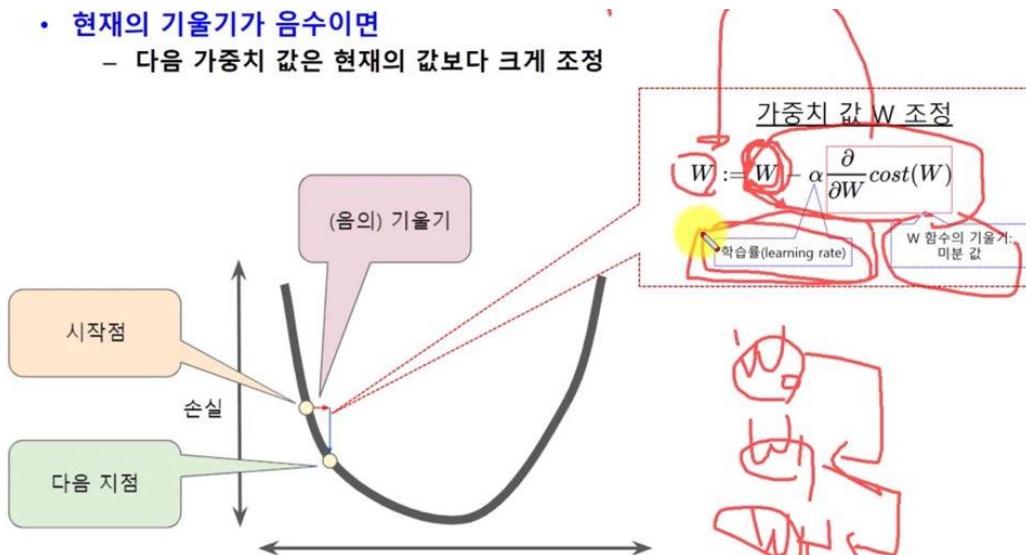


옵티마이저 발전과정



가중치

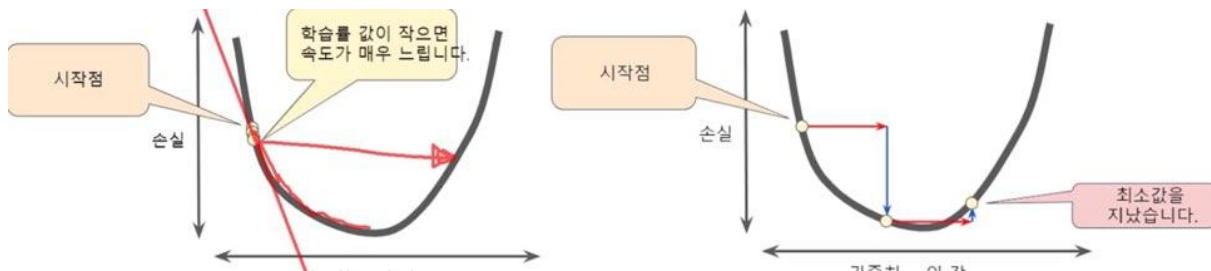
- 현재의 기울기가 음수이면
 - 다음 가중치 값은 현재의 값보다 크게 조정



학습률(learning rate) = 보폭

보폭을 너무 작게 하면 시간이 길어지고 크게 하면 이탈이 우려가 된다.

손실 함수 기울기가 작으면 더 큰 학습률이 시도 가능하다.



초매개변수

딥러닝에 우리가 설정하는 값 = '손잡이'

비용함수 = 손실함수 계산식 오차의 제곱

• 비용 함수(Cost function)

$$\text{cost}(W) = \frac{1}{n} \sum_i^n [y_i - H(x_i)]^2$$

• Cost를 최소화하는 W를 구하기 위한 식

- 해당 식은 접선의 기울기가 0이 될 때까지 반복

$$W := W - \alpha \frac{\partial}{\partial W} \text{cost}(W)$$

학습률(learning rate) W 함수의 기울기: 미분 값

비용함수로 최적의 w 구하는 식 = 평균제곱오차

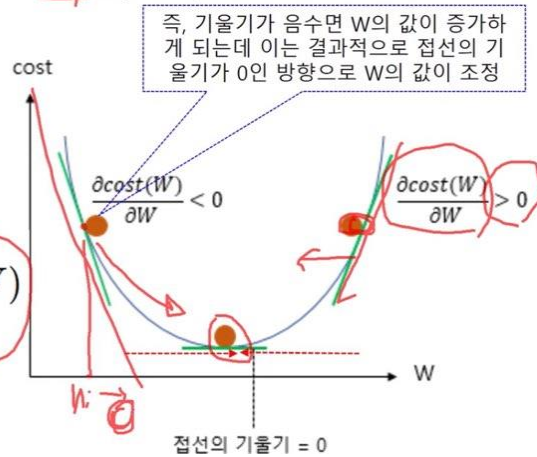
- 접선의 기울기가 음수일 때

$$W := W - \alpha(\text{음수}) = W + \alpha(\text{양수})$$

- 접선의 기울기가 양수일 때

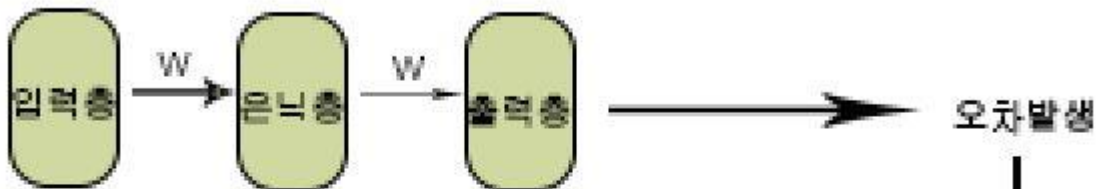
$$W := W - \alpha(\text{양수})$$

$$W := W - \alpha \frac{\partial}{\partial W} \text{cost}(W)$$

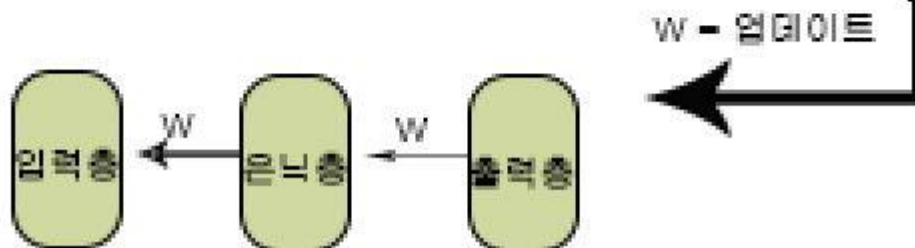


오차역전파

순전파 - 입력층 ~ 출력층 계산해 최종오차 계산 방법



역전파 - 오차 결과 값을 역으로 input 방향으로 다시보내며가중치 수정



속성과 데이터 소스 사이트(Attribute Information)

1. mpg: continuous
2. cylinders: multi-valued discrete
3. displacement: continuous
4. horsepower: continuous
5. weight: continuous
6. acceleration: continuous
7. model year: multi-valued discrete
8. origin: multi-valued discrete
9. car name: string (unique for each instance): 없음

이진(이항) 분류

두가지 분류 방법

- PASS / FAIL
- SPAM / HAM
- Positive & Negative
- 로지스틱 회귀라고 부른다.

결과 기술 방식

- 4 개의 결과(일반 레이블 방식 & One Hot Encoding 방식)