

Idea Factory Intensive Program #2

딥러닝 홀로서기

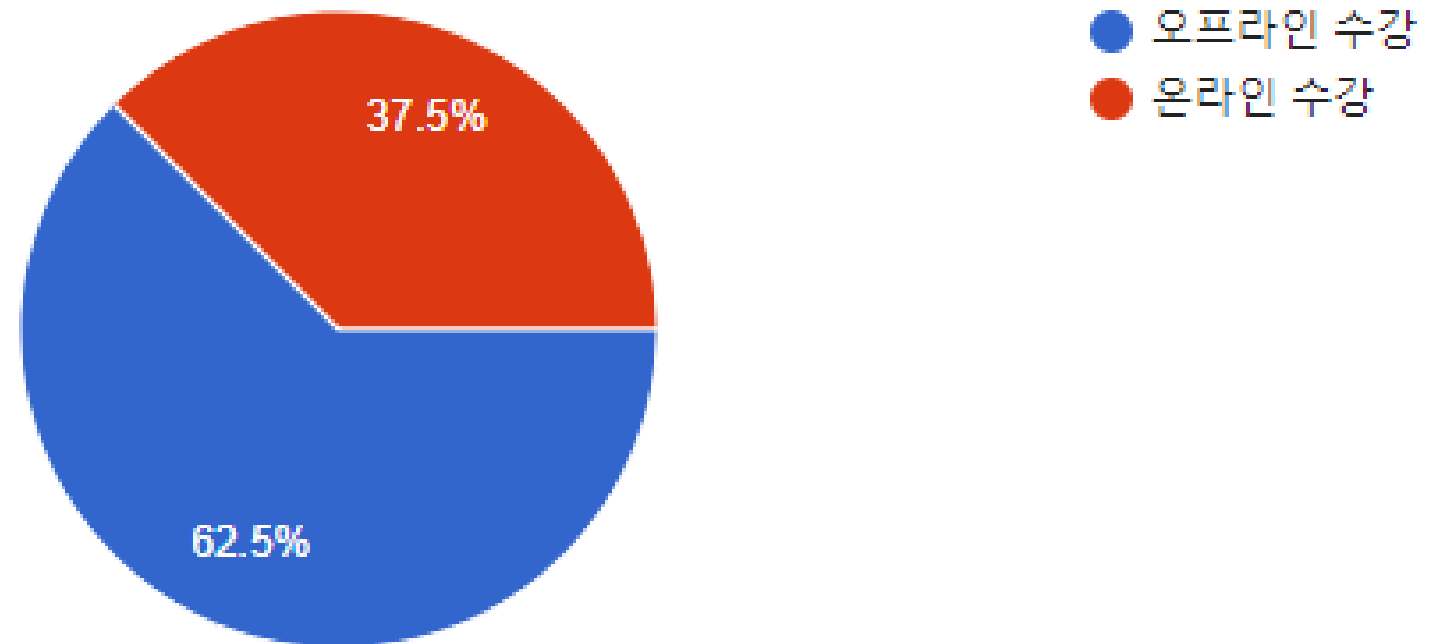
#2

이론강의/PyTorch실습/코드리뷰

딥러닝(Deep Learning)에 관심이 있는 학생 발굴을 통한
딥러닝의 이론적 배경 강의 및 오픈소스 딥러닝 라이브러리 PyTorch를 활용한 실습

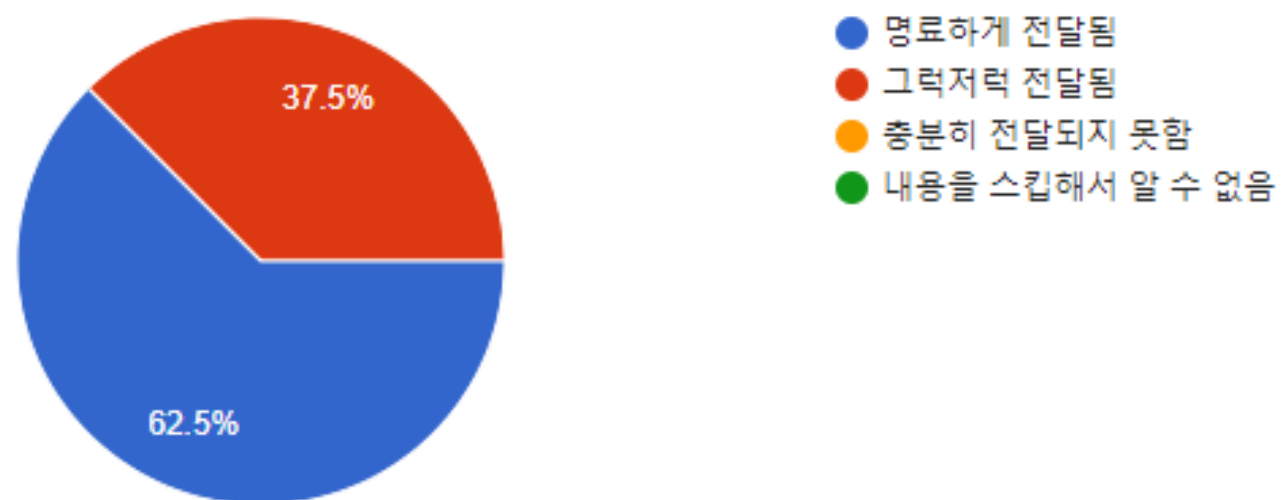
어떤 방식을 통해 강의를 접하셨나요?

응답 8개



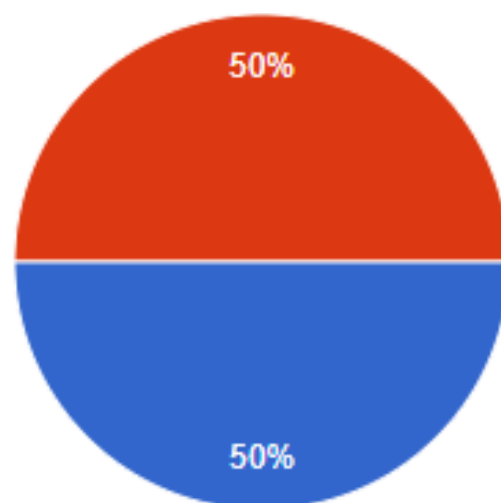
이번 수업에서 본 세미나의 목적이나 앞으로의 방향, 딥러닝 공부의 중요성이 잘 전달되었을까요?

응답 8개



이번 수업에서 Machine Learning Basic에 대한 내용이 잘 전달되었을까요?

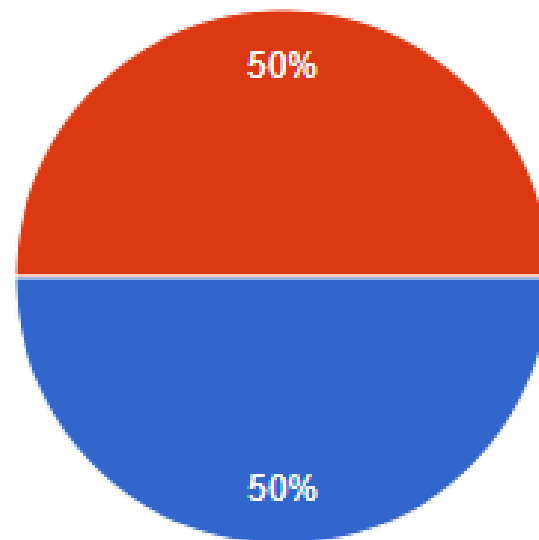
응답 8개



- 명료하게 전달됨
- 그럭저럭 전달됨
- 충분히 전달되지 못함
- 내용을 스킵해서 알 수 없음

이번 수업에서 Linear Regression에 대한 내용이 잘 전달되었을까요?

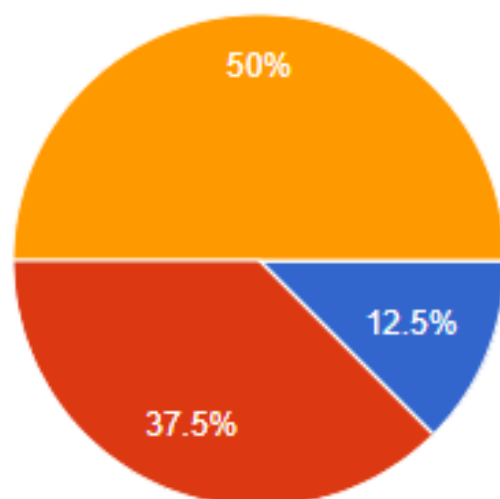
응답 8개



- 명료하게 전달됨
- 그럭저럭 전달됨
- 충분히 전달되지 못함
- 내용을 스킵해서 알 수 없음

이번 수업에서 Linear Regression 코딩 실습에 대한 내용이 잘 전달되었을까요?

응답 8개



- 명료하게 전달됨
- 그럭저럭 전달됨
- 충분히 전달되지 못함
- 내용을 스킵해서 알 수 없음

Feedback from Last Lecture

다양한 지식적 기반을 가진 학생들이 있는데, 앞으로의 수업에서 어떤 학생들을 타겟으로 할 것인가?

→ Python과 Numpy를 어느 정도 다룰 줄 아는 학생들을 타겟으로 하며, 이러한 기본 지식에 대한 참고 자료들을 앞으로도 공지 예정.

주차별 계획은?

→ 포스터에 나와 있는 계획을 예정하였으나, 이는 수업 과정에서 변경될 수 있음.

기본적으로 딥러닝에 대한 전반적 이해와 다양한 모델을 소개하는 정도로 마무리 될 예정.

각 수업 시작 때 마다 그날의 예상 진도를 공지해 줄 예정.

실습에 관한 의견들

→ 기존 튜토리얼들이 이미 짜여진 코드들을 설명하는데 그쳤다면, 저번 수업에서는 스크래치부터 짜는 모습을 보여주었다. 이에 대해 점진적 과정을 보는 것과, 문제가 생겼을 때 해결하는 방법을 볼 수 있는 점이 좋았으나, 코드가 미리 준비되어 있는 것이 좋을 것 같다는 의견을 받았음.

앞으로는 실습에 사용될 완성된 코드를 미리 준비하여 Github에서 미리 받아볼 수 있도록 할 것이며, 실습 시간에는 그 코드를 다시 코딩해 나가는 과정을 보여줄 예정.

Topics to learn today

1. Review from last lecture

Problems of ML / Linear Regression

Linear Regression with Pytorch

2. Binary/Multinomial Classification Problem

with Logistic Regression

Multinomial Classification with Pytorch

3. History of Deep Learning

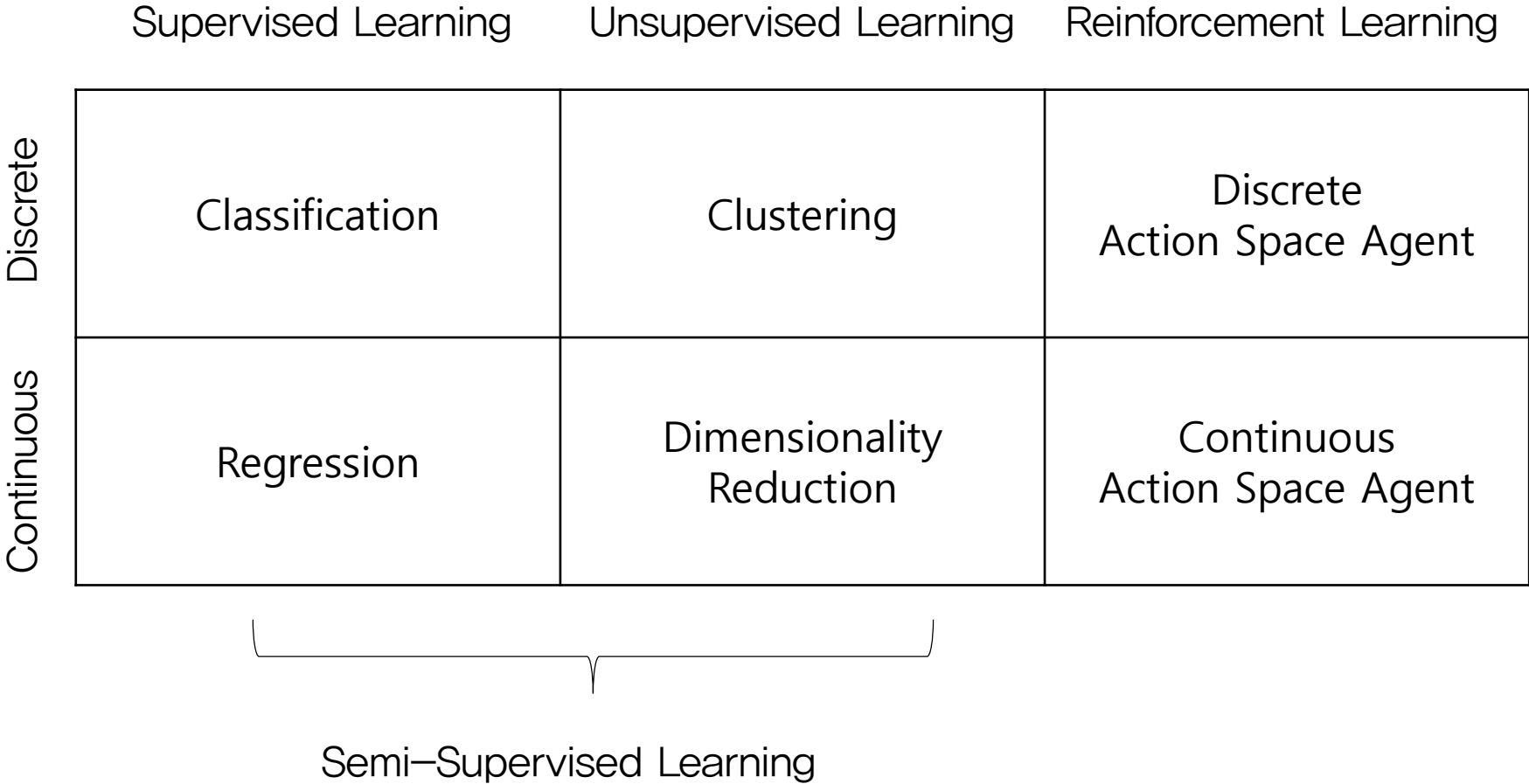
from simple perceptron to CNN

4. Solving XOR Problem with MLP

Feed forward / Backpropagation

Solving Regression and Classification Problem with MLP

Review from Last Lecture



Review from Last Lecture

Hypothesis

Model

Cost

Loss

Optimization

Review from Last Lecture

Hypothesis

Model

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

Cost

Loss

$$cost(W, b) = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$

Optimization

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

Review from Last Lecture

Hypothesis

Model

$$H(X) = XW$$

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11}w_1 + x_{12}w_2 + x_{13}w_3 \\ x_{21}w_1 + x_{22}w_2 + x_{23}w_3 \\ x_{31}w_1 + x_{32}w_2 + x_{33}w_3 \\ x_{41}w_1 + x_{42}w_2 + x_{43}w_3 \\ x_{51}w_1 + x_{52}w_2 + x_{53}w_3 \end{pmatrix}$$

Cost

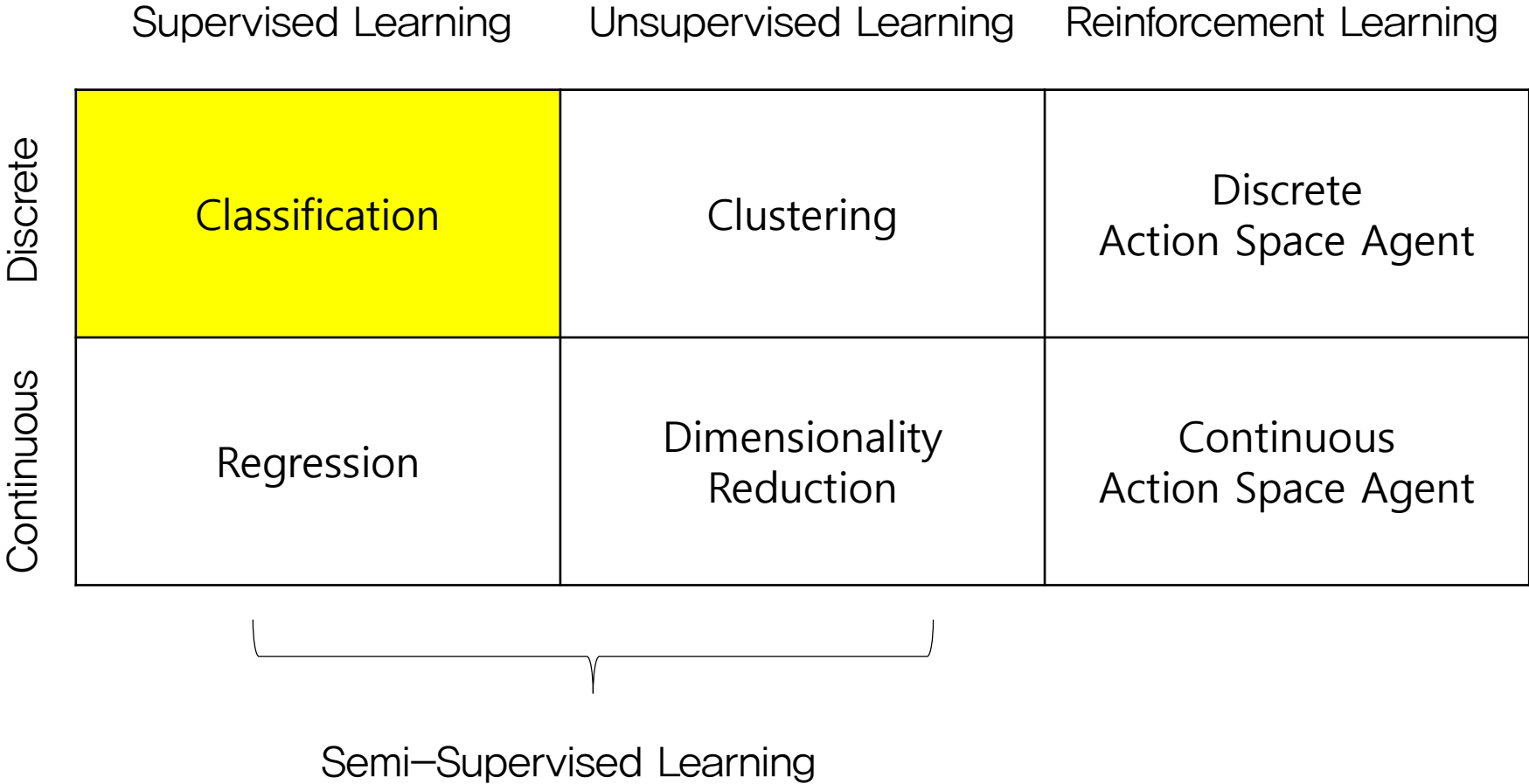
Loss

$$cost(W, b) = \frac{1}{m} \sum_{i=1}^m (H(x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}) - y^{(i)})^2$$

Optimization

Binary Classification

Binary Classification



Example of Binary Classification



Dog or Cat?



Show this
facebook feed to
user or not?

0, 1 encoding

Dog(0), Cat(1)

Show the feed (0)
Do not show the feed (1)

Binary Classification Hypothesis

Pass(1)/Fail(0) based on study hours

Can we use Linear Regression?

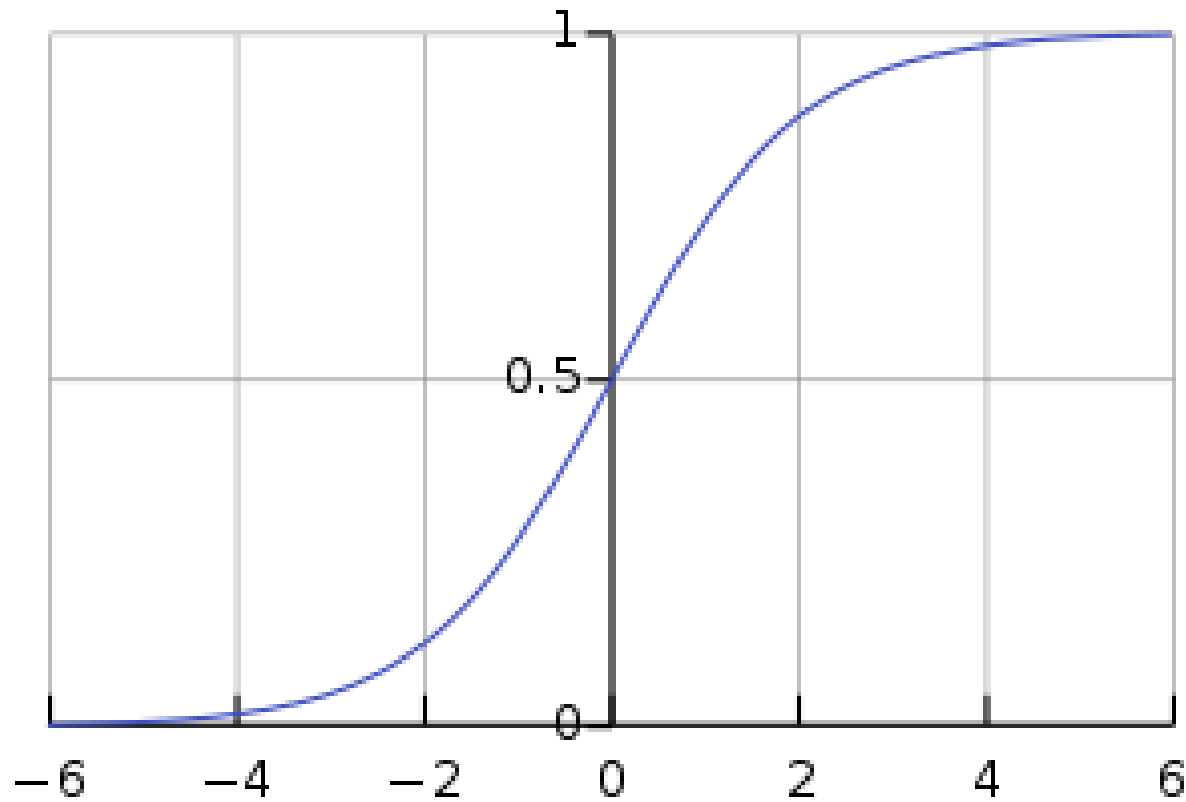


Logistic Hypothesis

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

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

Sigmoid (Logistic) Function



$$G(z) = \frac{1}{1 + e^{-z}}$$

Logistic Hypothesis

$$H(X) = \frac{1}{1 + e^{WX}}$$

Binary Classification Cost

Cost Function

Can we use the cost function of linear regression?

$$cost(W, b) = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2 \qquad H(X) = \frac{1}{1 + e^{WX}}$$

Cross Entropy

Difference between two probability distribution

$$H(P, Q) = - \sum P(x) \log(Q(x))$$

Cost Function

$$\textit{cost}(W) = \frac{1}{m} \sum_{i=1}^n c(H(x^{(i)}), y^{(i)})$$

$$c(H(x), y) = \begin{cases} -\log(H(x)) & : y = 1 \\ -\log(1 - H(x)) & : y = 0 \end{cases}$$

Cost Function

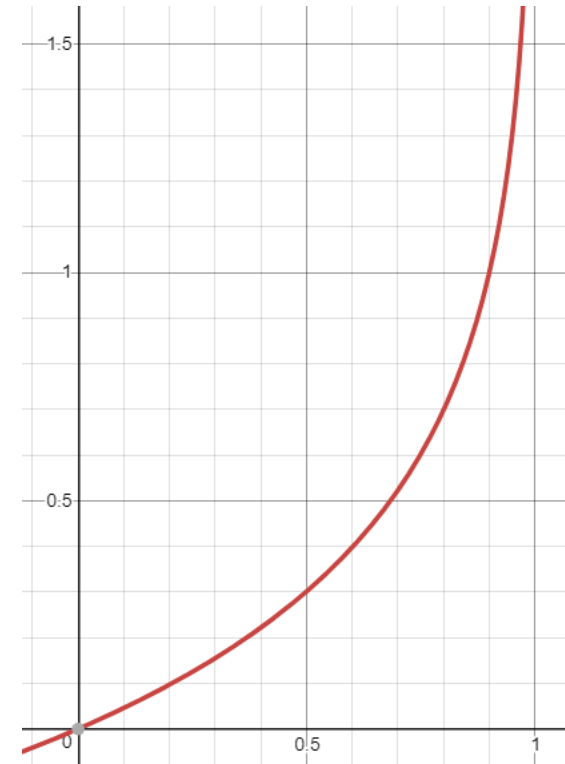
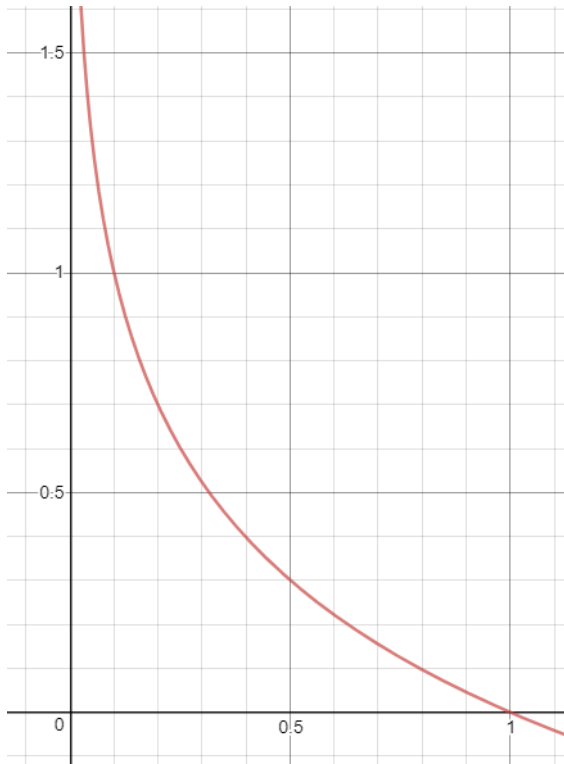
$$\text{cost}(W) = \frac{1}{m} \sum_{i=1}^n c(H(x^{(i)}), y^{(i)})$$

$$c(H(x), y) = \begin{cases} -\log(H(x)) & : y = 1 \\ -\log(1 - H(x)) & : y = 0 \end{cases}$$

$$c(H(x), y) = -y \log(H(x)) - (1 - y) \log(1 - H(x))$$

Cost Function

$$c(H(x), y) = \begin{cases} -\log(H(x)) & : y = 1 \\ -\log(1 - H(x)) & : y = 0 \end{cases}$$



Multinomial Classification

Binary Classification

Multinomial Classification

Multinomial Classification Hypothesis

Multinomial Classification

$$H(X) = XW + b$$

Multinomial Classification

$$H(X) = G(XW + b)$$

Softmax function

$$\textit{Softmax}(\hat{y}_i) = \frac{e^{\hat{y}_i}}{\sum_j e^{\hat{y}_j}}$$

Softmax function

Multinomial Classification Cost

Cost function

S(y)

y

0.7

1

0.2

0

0.1

0

$$Cost(W, b) = \sum_i y_i \log(S(WX + b))$$

$$\hat{y} = WX + b$$

Binary and Multinomial Classification

Binary

Multinomial

Hypothesis

$$\textit{Sigmoid}(WX)$$

$$\textit{Softmax}(WX)$$

Cost

$$-y\log(H(x)) - (1 - y)\log(1 - H(x))$$

$$\sum -y\log(H(X))$$