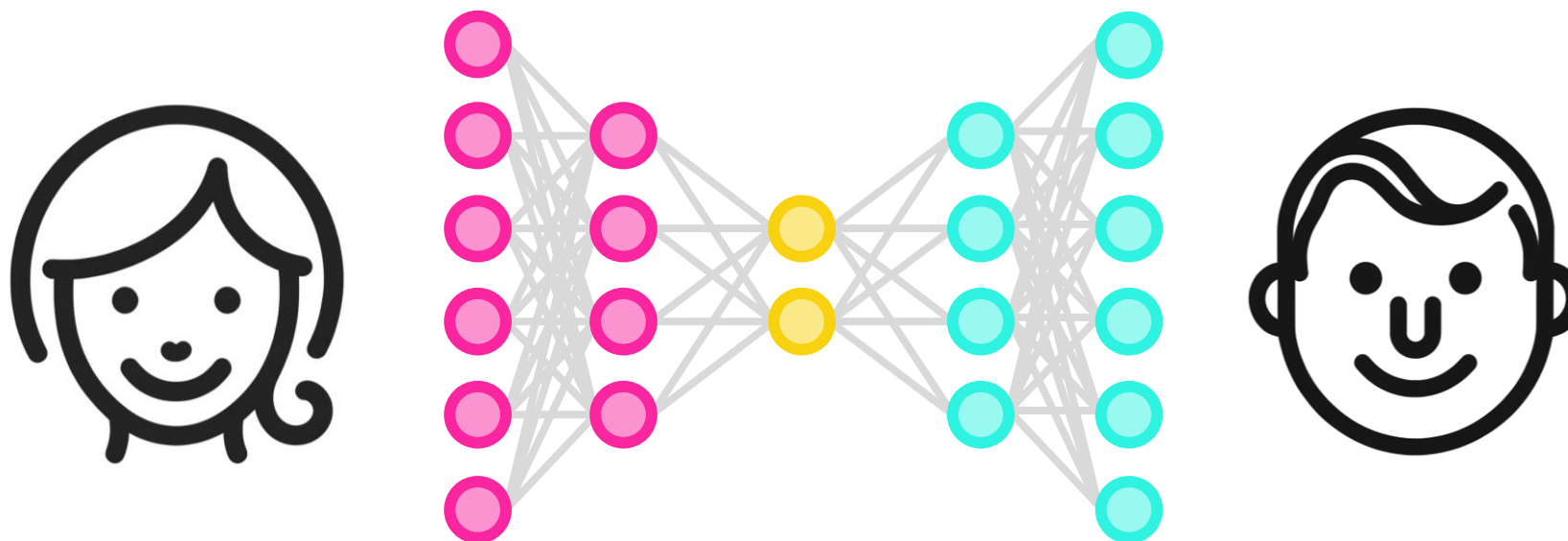


Deep Learning 101

오토인코더

Autoencoder



안녕하세요 신박AI입니다

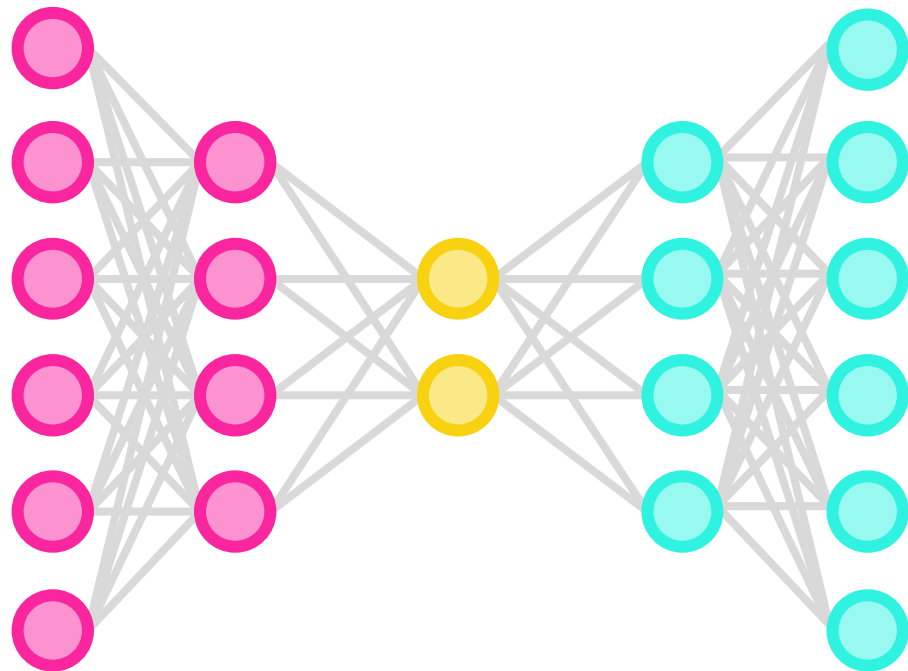


오늘 여러분께 소개해드리고 싶은
모델은 바로 오토인코더입니다.

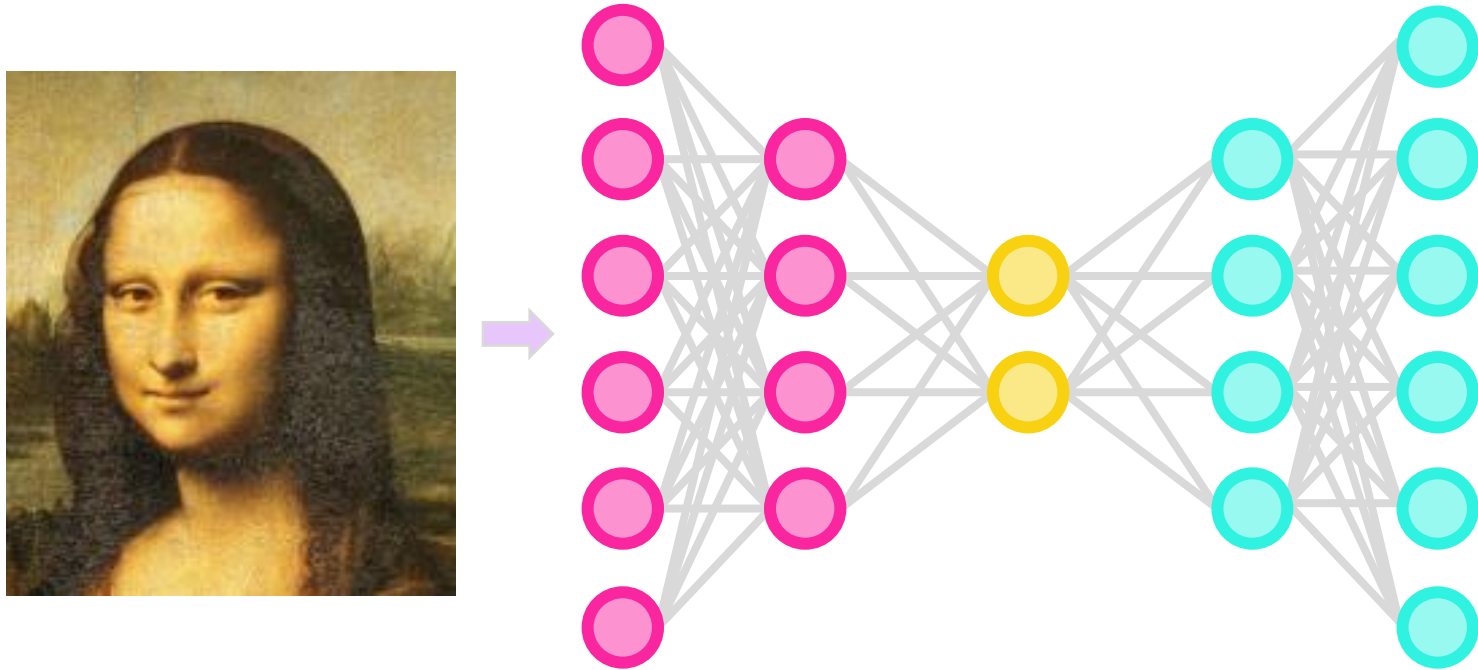
오토인코더는 구조도 간단하고
활용할 영역도 많은
좋은 신경망입니다

오늘 영상에서는 오토인코더의
중요 개념, 학습 알고리즘, 그리고
활용 영역에 대해 간략하게 다루어
보겠습니다.

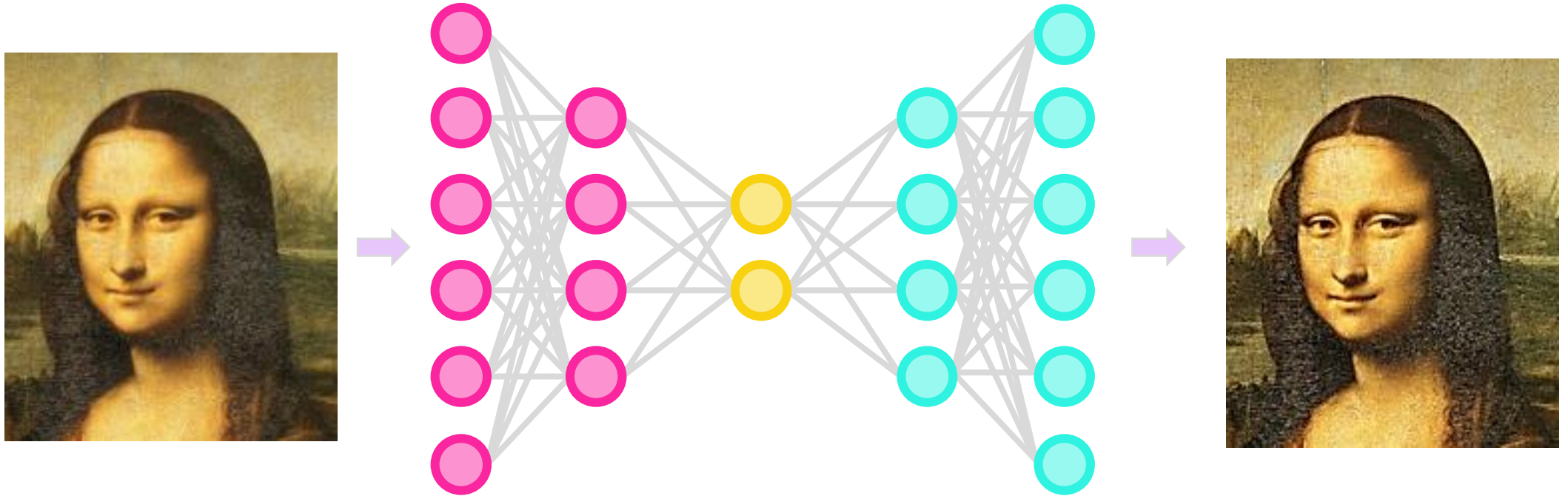
오토인코더는 데이터의 효율적인 표현을 학습하는 인공 신경망 모델입니다.



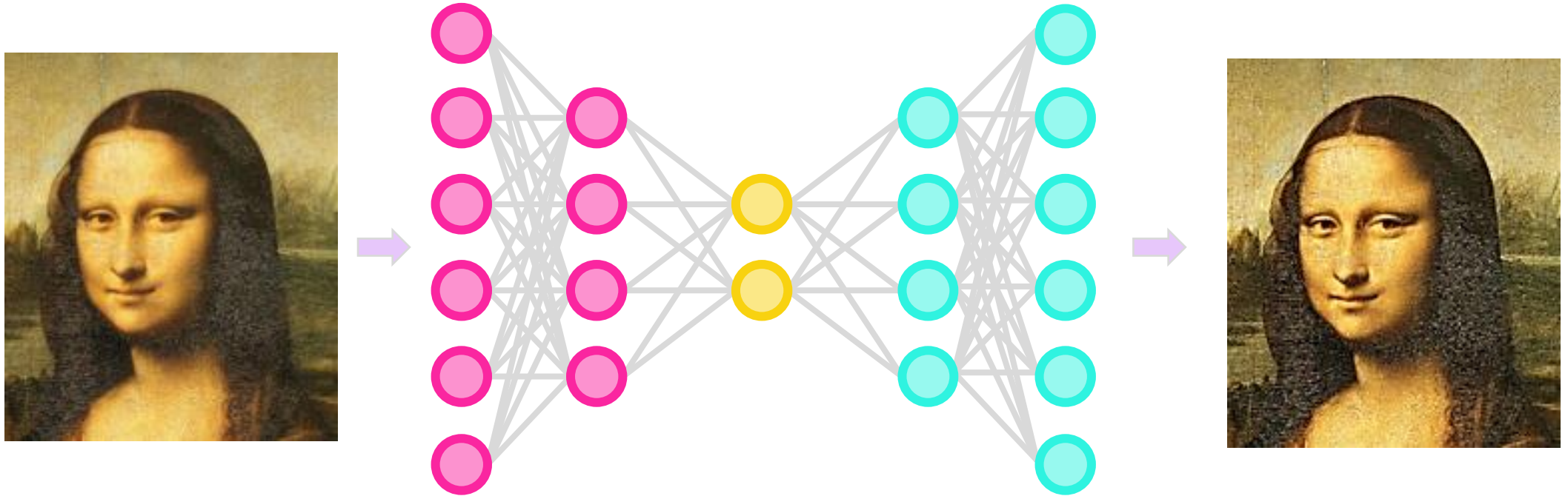
이 모델은 데이터의 중요한 특성을 추출하고,



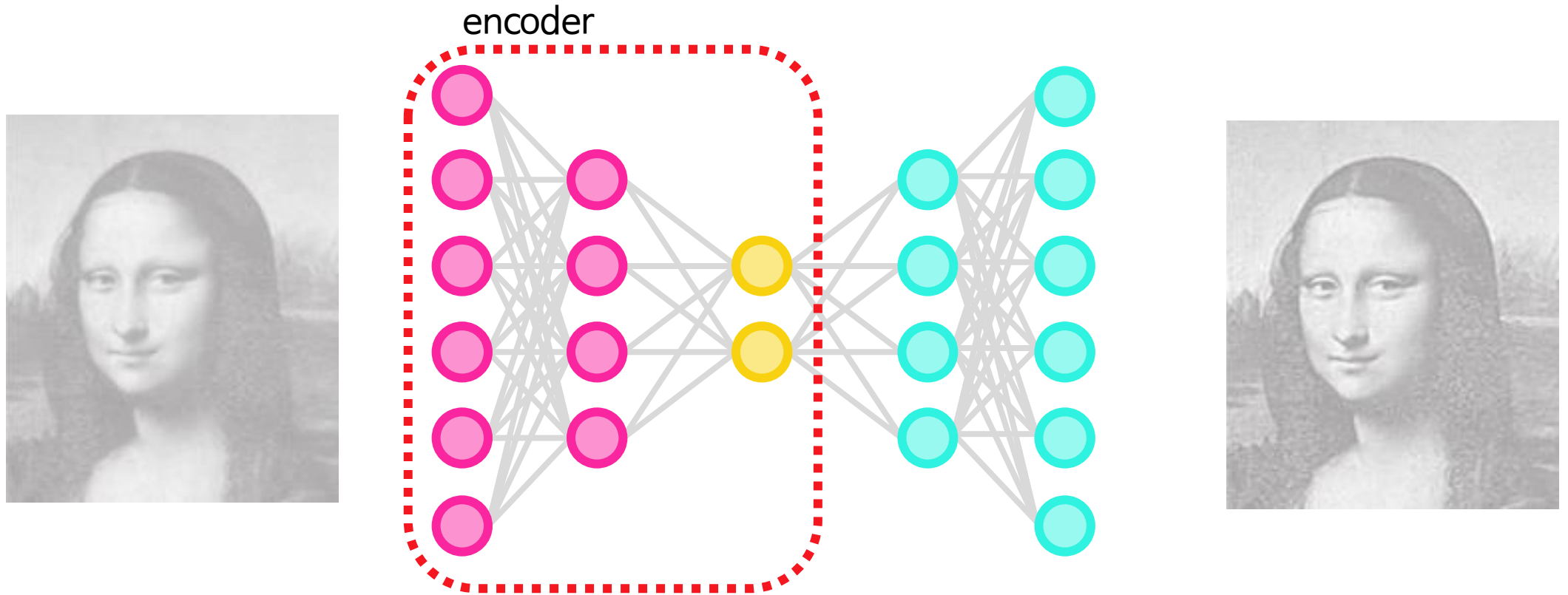
이를 바탕으로 데이터를 압축한 후 다시 복원하는 과정을 거치게 됩니다.



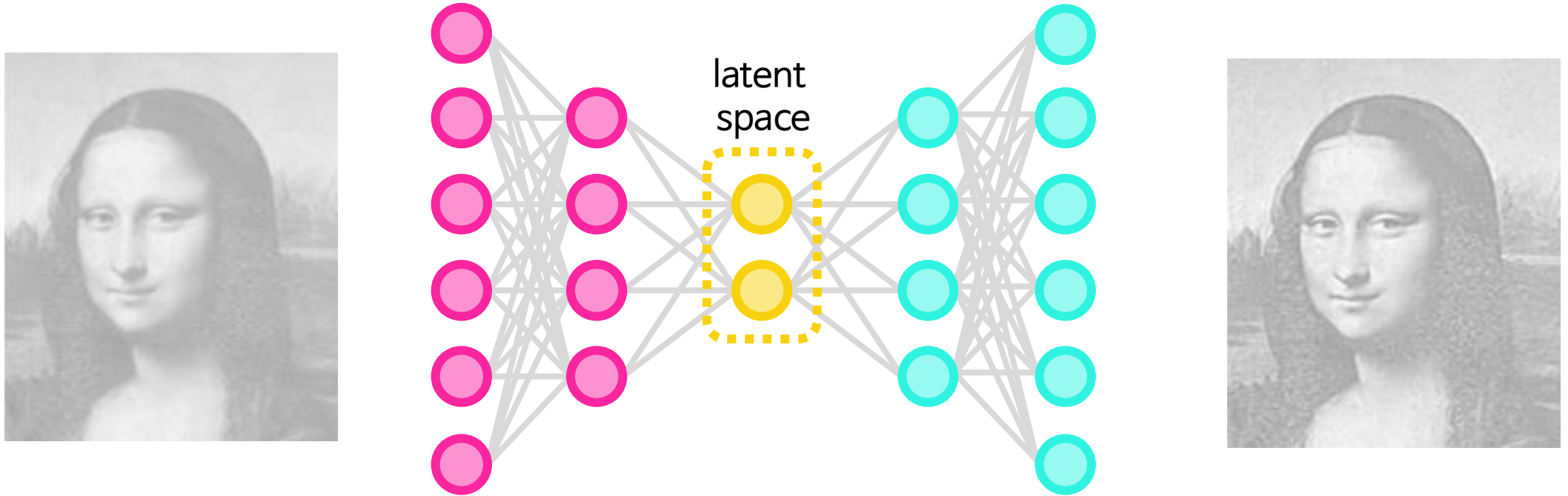
오토인코더는 크게 두 부분으로 구성되어 있습니다.



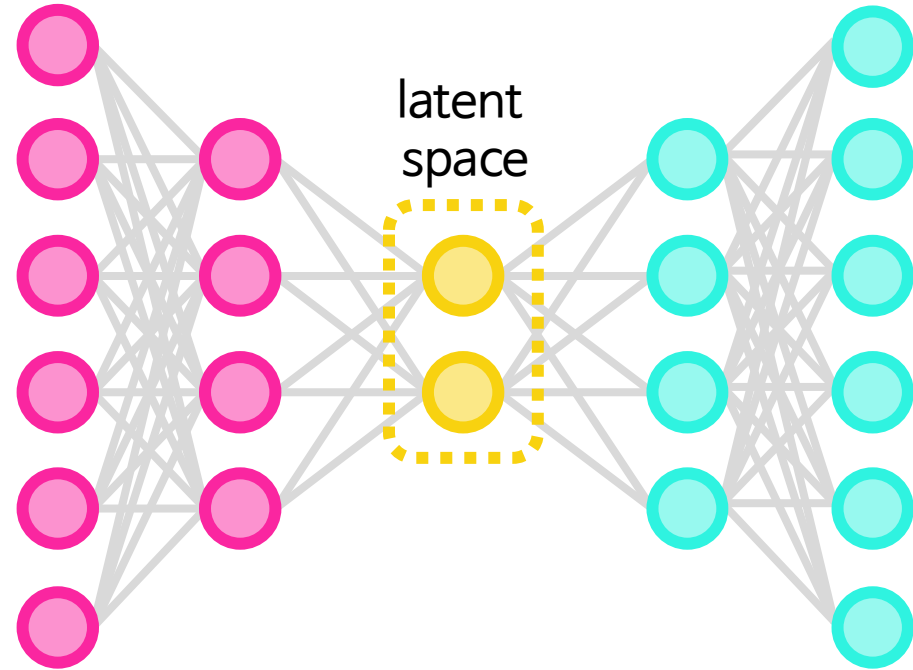
이 부분은 인코더라고 합니다. 입력 데이터를 받아서, 그것을 더 작은 차원으로 압축합니다.



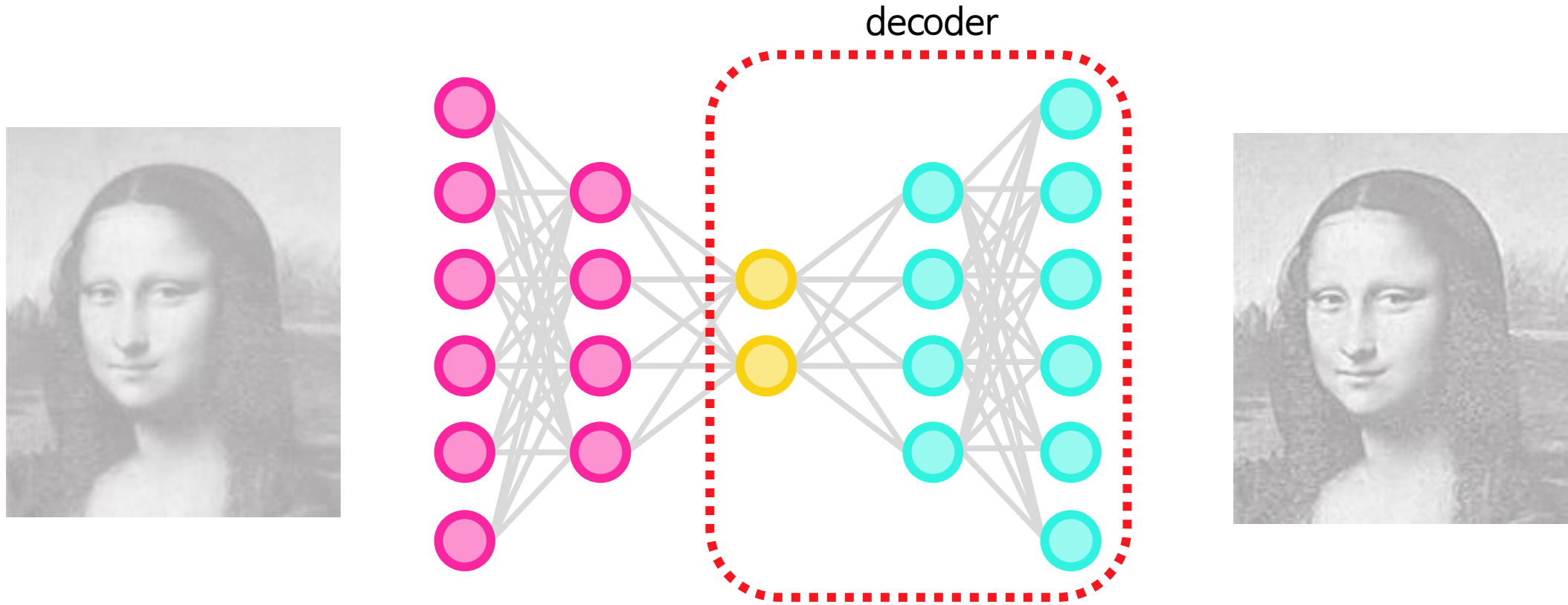
이 압축된 표현을 ‘잠재 공간(latent space)’ 또는 ‘코딩(coding)’이라고 합니다.



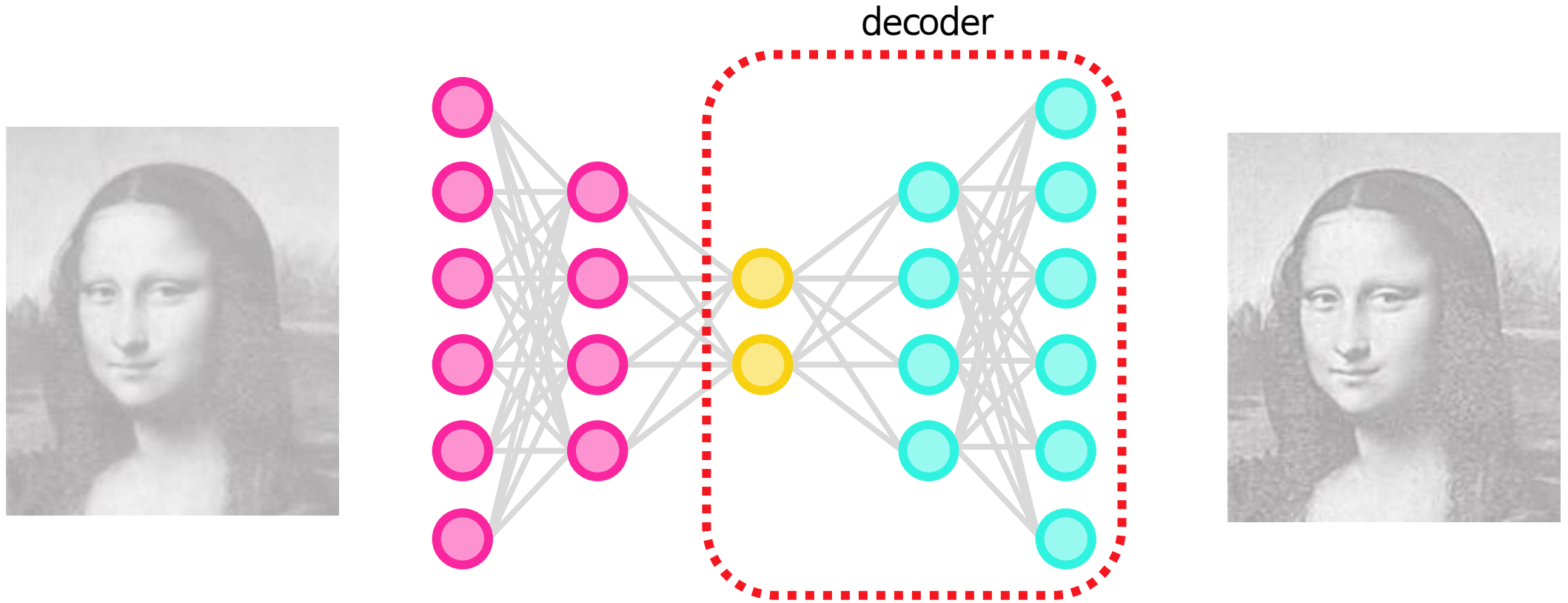
여기서 중요한 것은 데이터의 중요한 특징을 얼마나 잘 잡아내느냐
입니다.



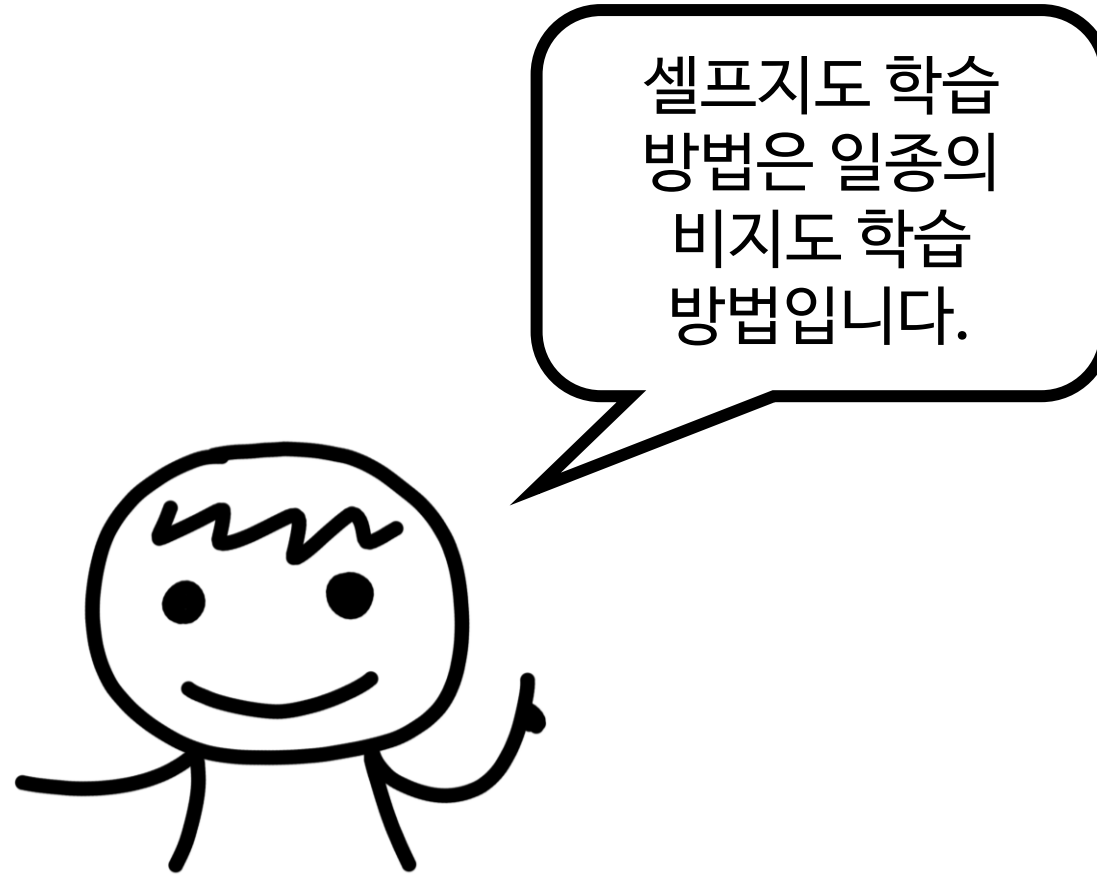
이 부분은 디코더라고 부르는데요, 디코더는 인코더에서 만들어진 압축된 표현을 다시 원래의 데이터로 복원하려 합니다.



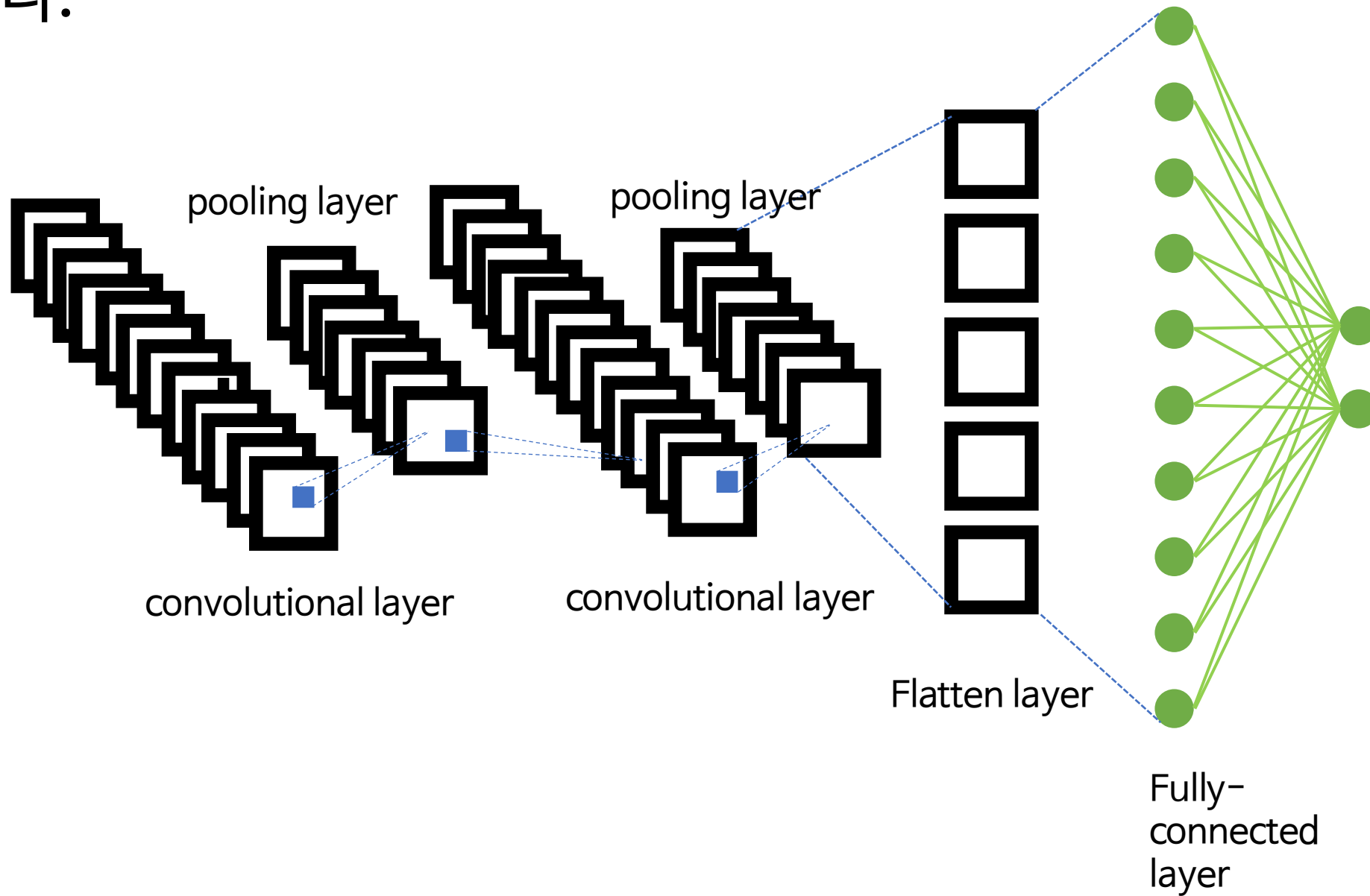
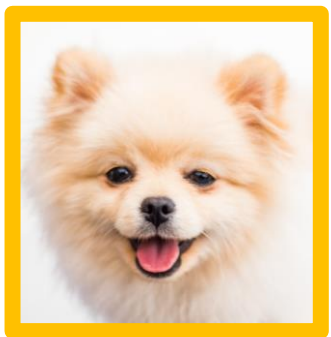
완벽하게 복원하는 것은 아니지만, 가능한 한 원본에 가깝게 복원합니다.



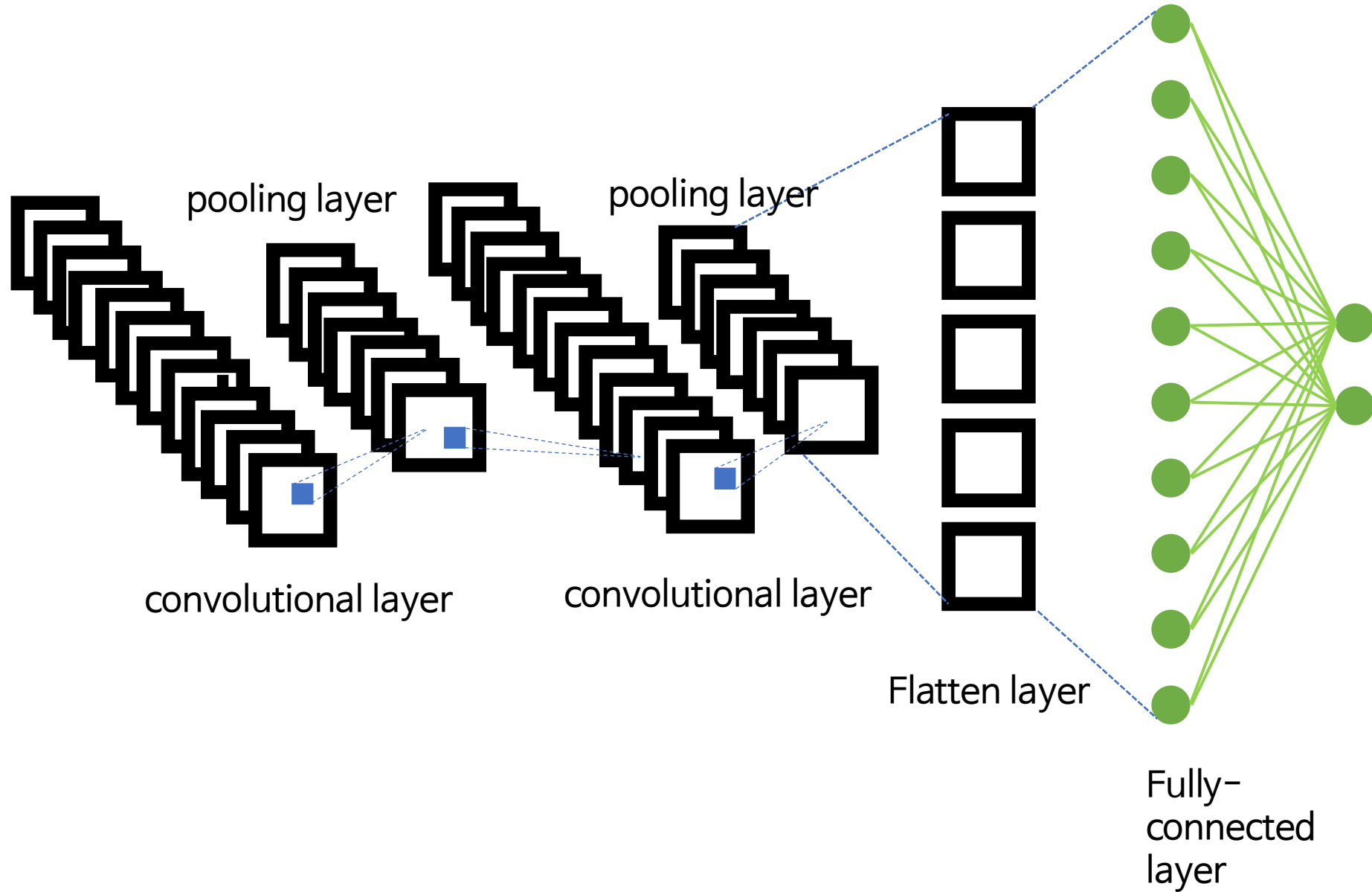
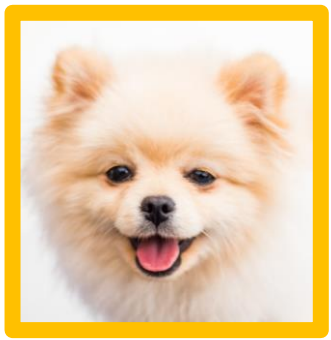
오토인코더의 학습은 셀프지도 학습 방식으로 이루어집니다.



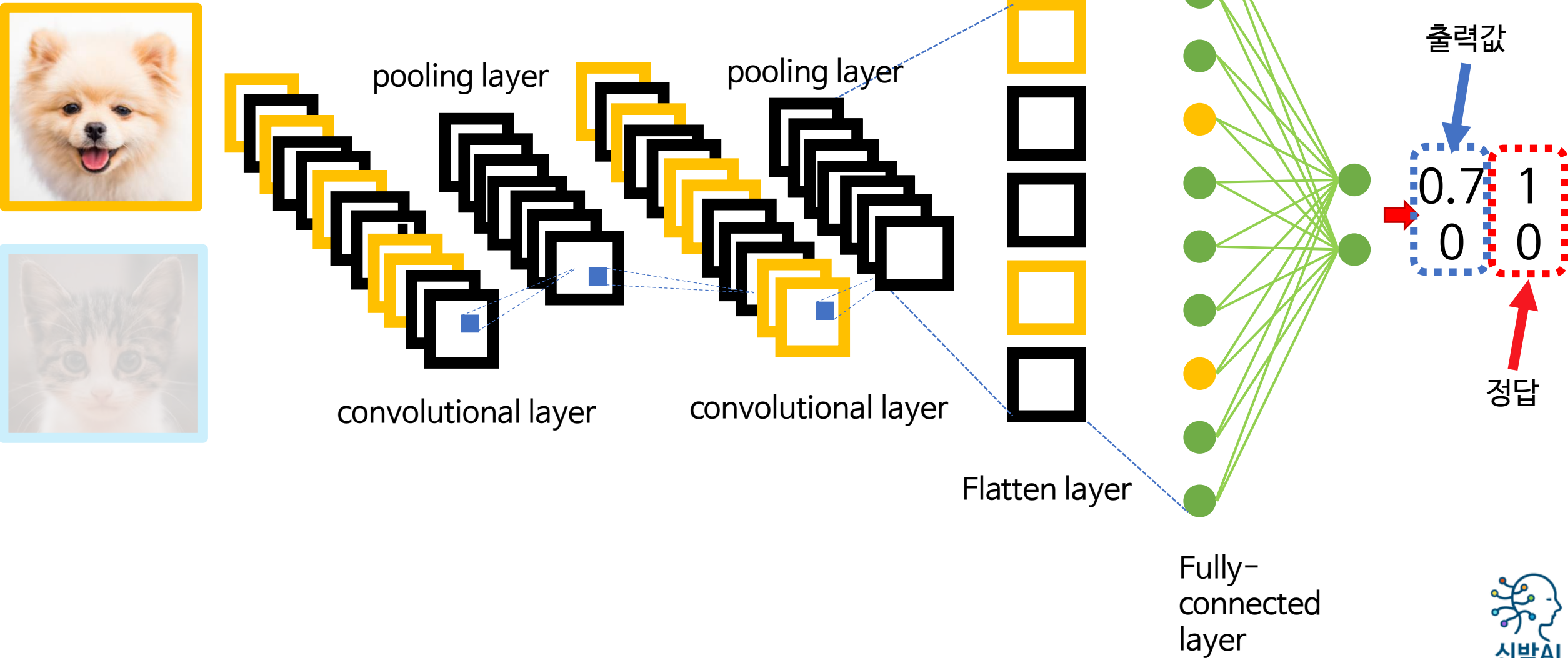
지금까지 우리는 주로 지도학습법, 교사학습법으로 신경망을 학습해왔습니다.



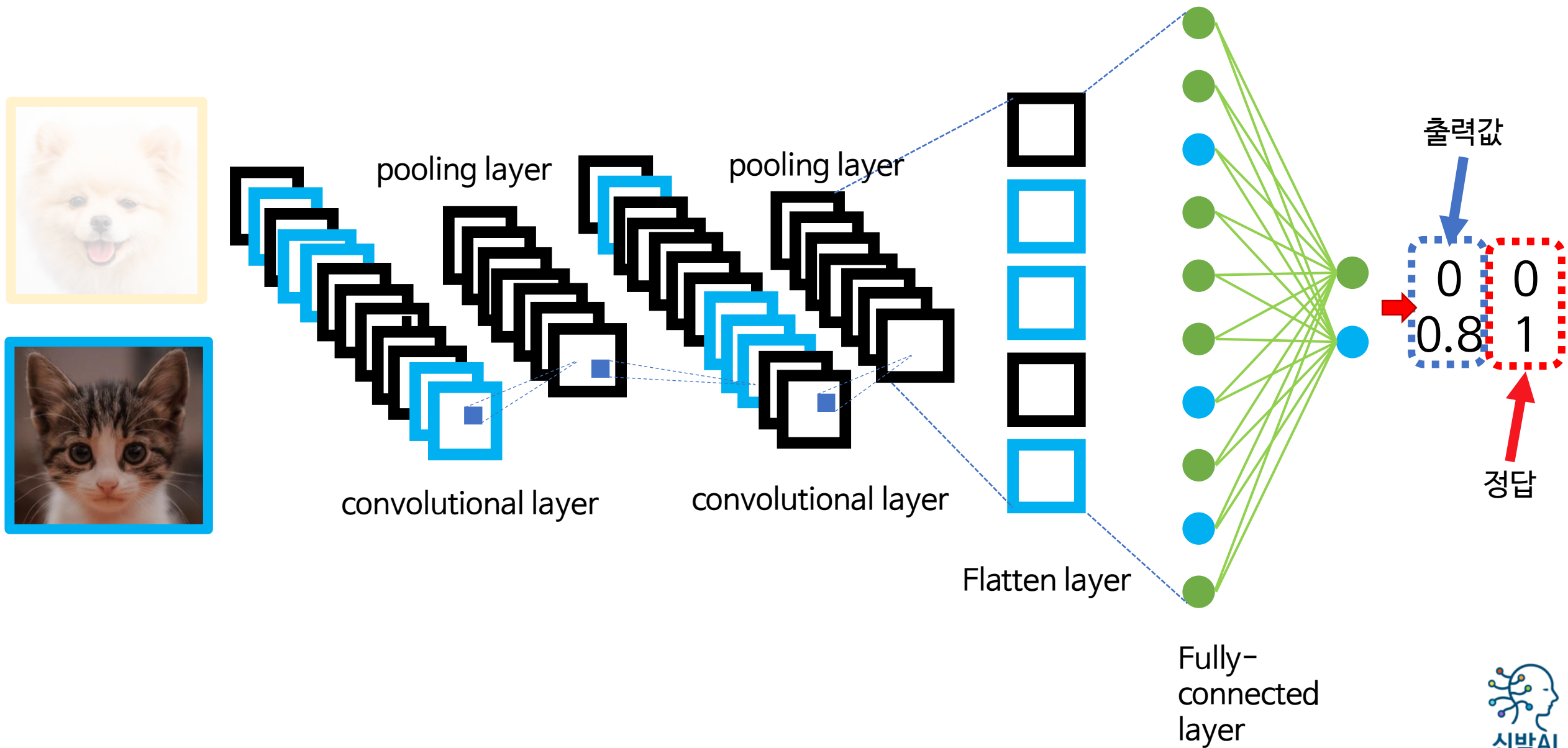
예를들면 개와 고양이를 분류하는 CNN의 경우는,



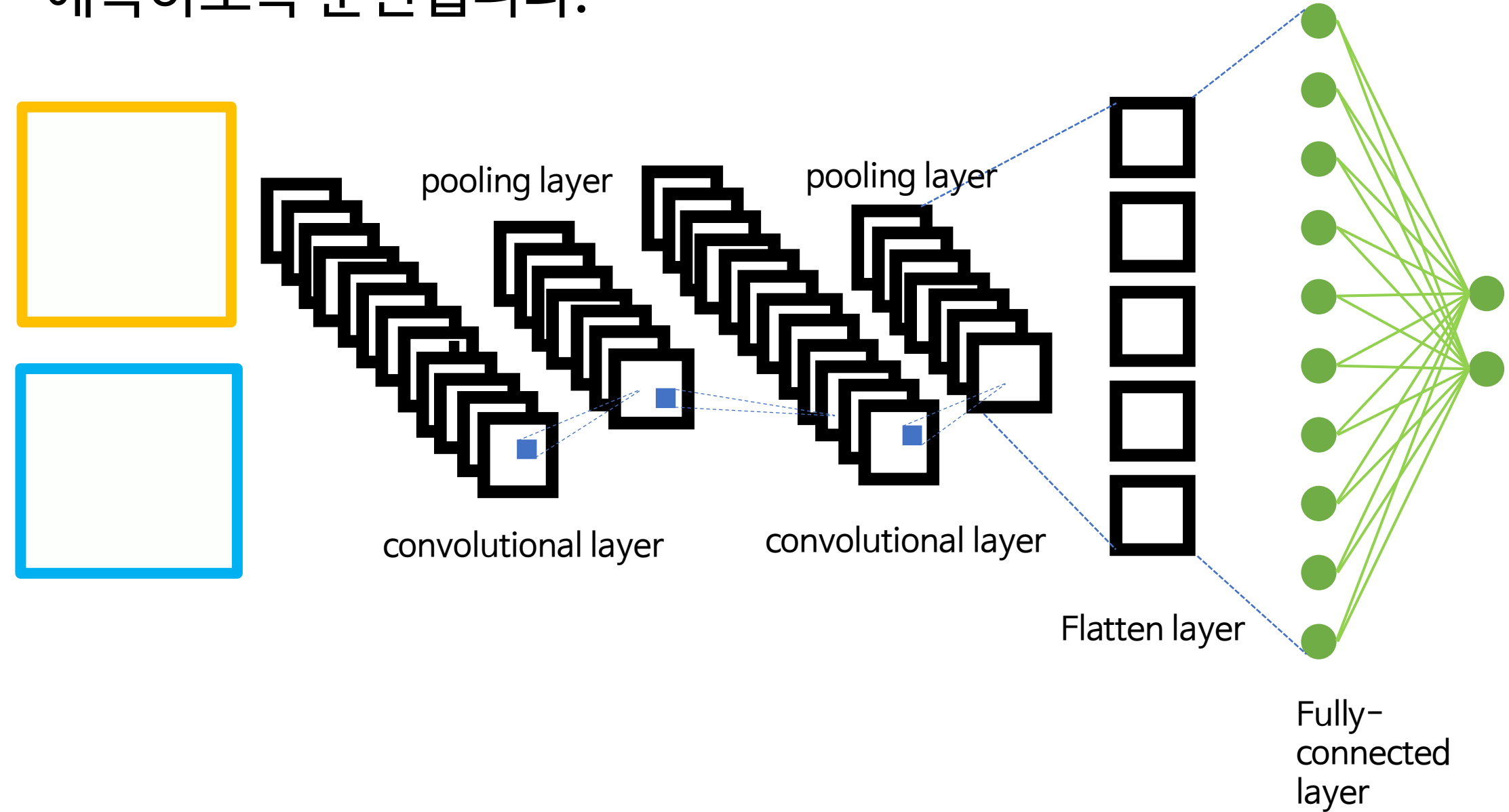
주어진 이미지셋에 대응하는 정답 (라벨, label)을 주어서,



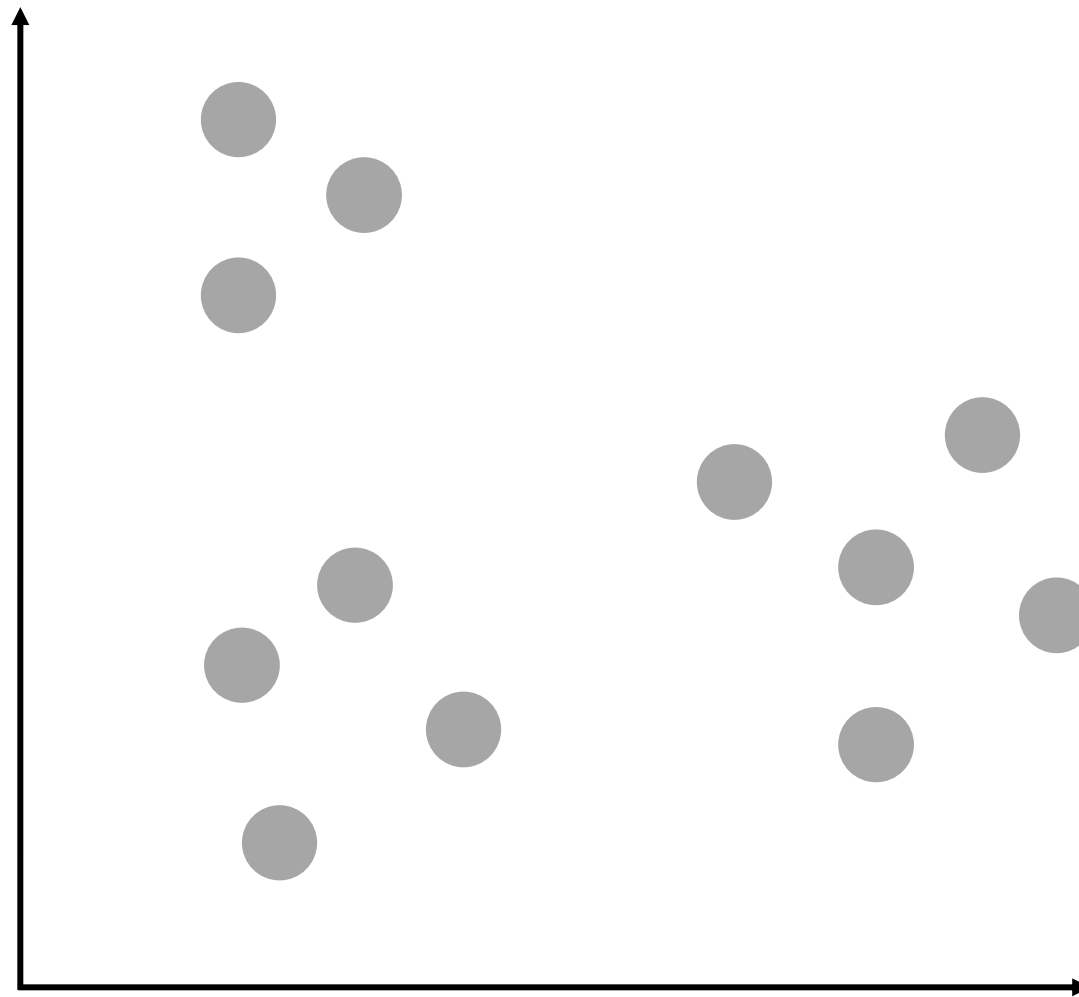
이미지의 특징과 정답간의 관계성을 신경망이 학습하는 방식입니다.



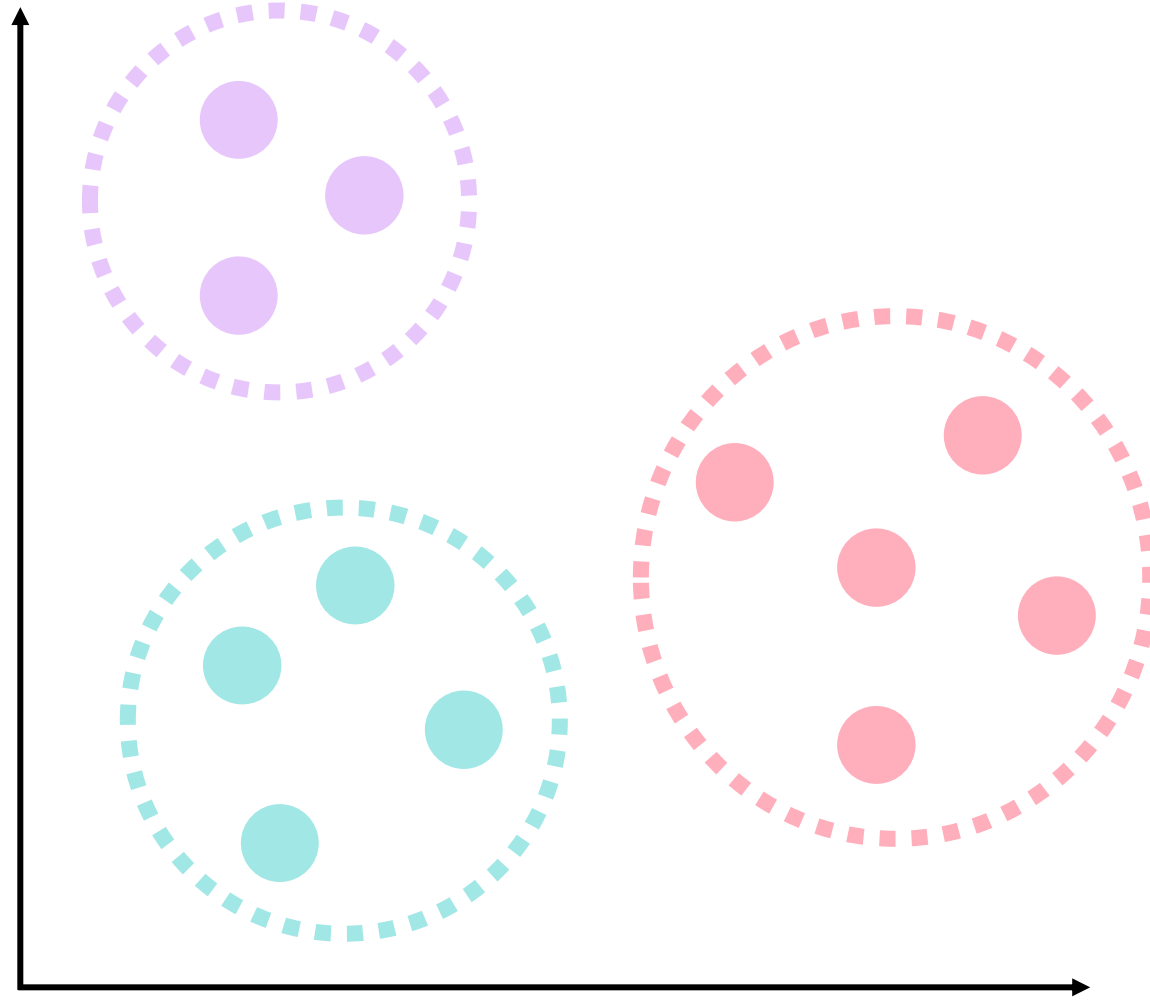
지도 학습방법은 신경망이 주어진 입력 데이터에 대해 정확한 출력을 예측하도록 훈련됩니다.



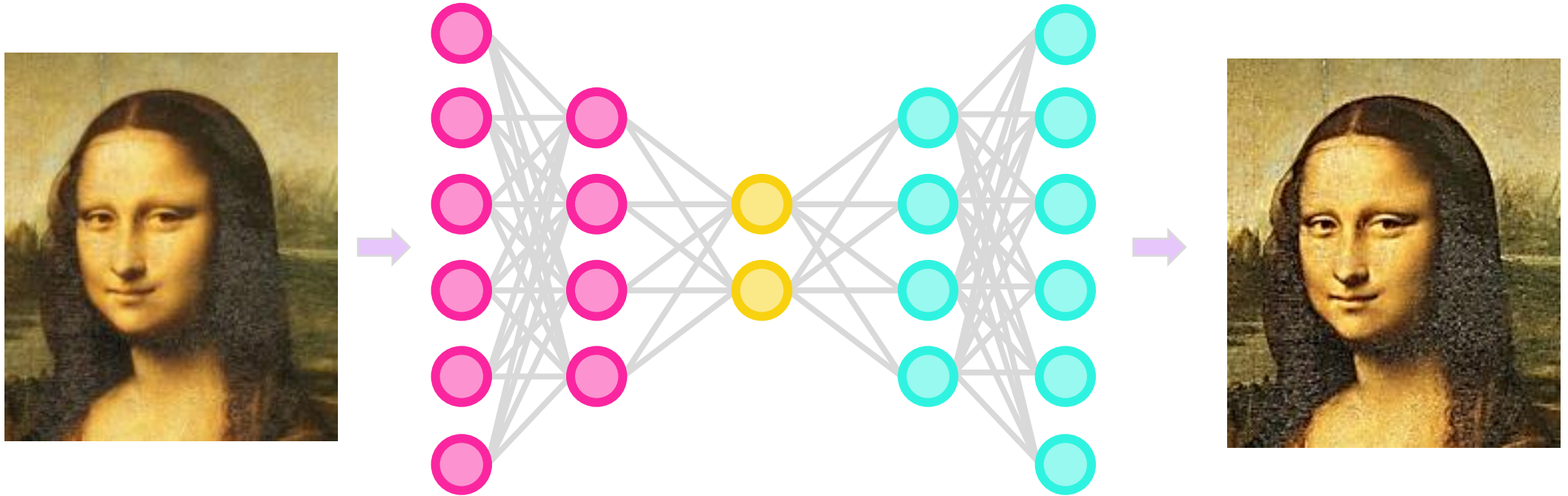
그러나 비지도 학습방법은 라벨이 없는 데이터를 사용하여 모델이 데이터에 내재된 구조나 패턴을 스스로 찾아내도록 하는 방식입니다.



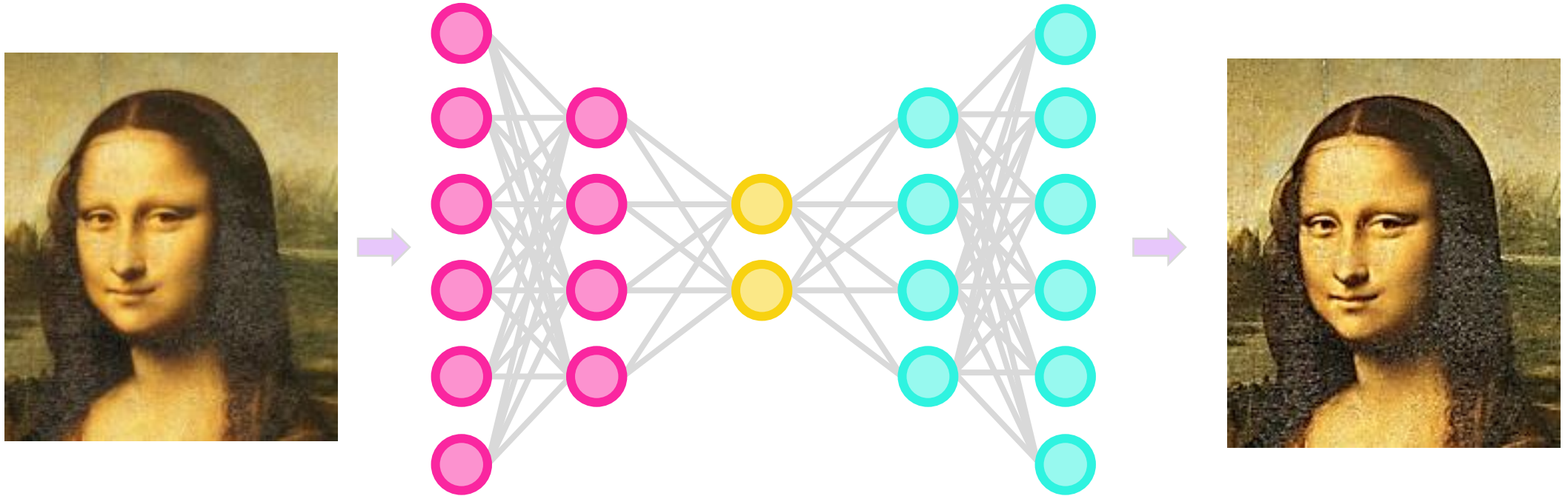
그러나 비지도 학습방법은 라벨이 없는 데이터를 사용하여 모델이 데이터에 내재된 구조나 패턴을 스스로 찾아내도록 하는 방식입니다.



오토인코더의 학습에도 출력값(라벨)이 주어지지 않으며,



오토인코더는 주어진 데이터의 특성과 분포를 기반으로 학습합니다.



이는 레이블이 없는 데이터만을 사용하여, 원본 데이터와 복원된 데이터 간의 차이,



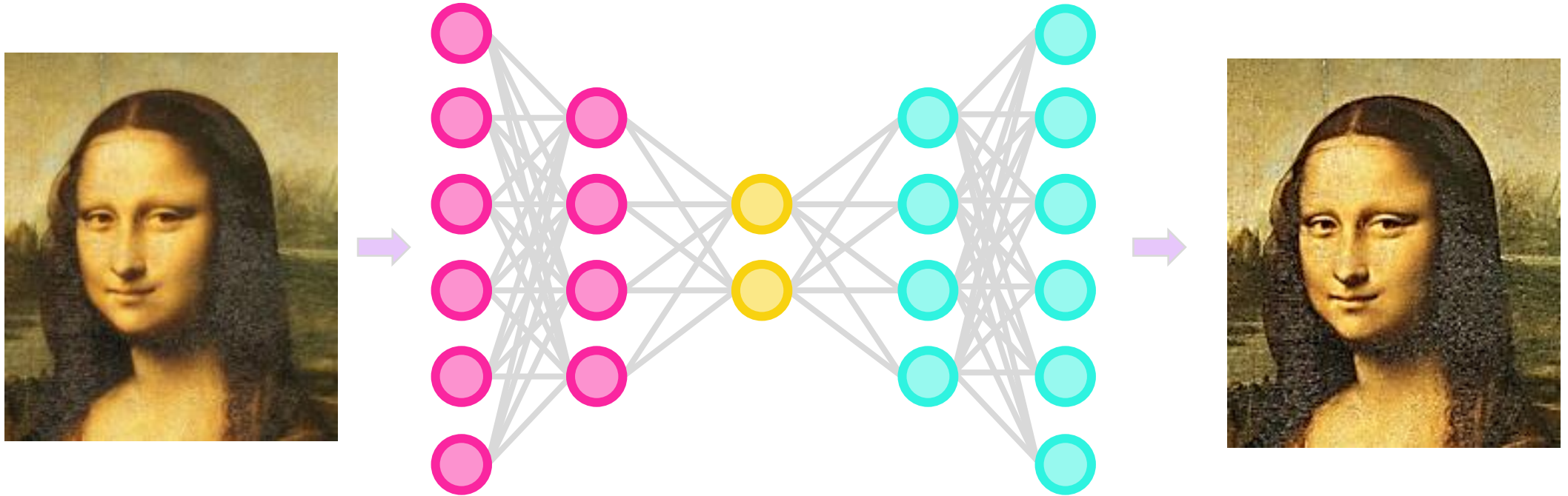
즉 재구성 손실 reconstruction loss 을 최소화하는 방향으로
진행됩니다.

$$\left(\text{Original Image} - \text{Reconstructed Image} \right) = \text{Reconstruction Loss}$$

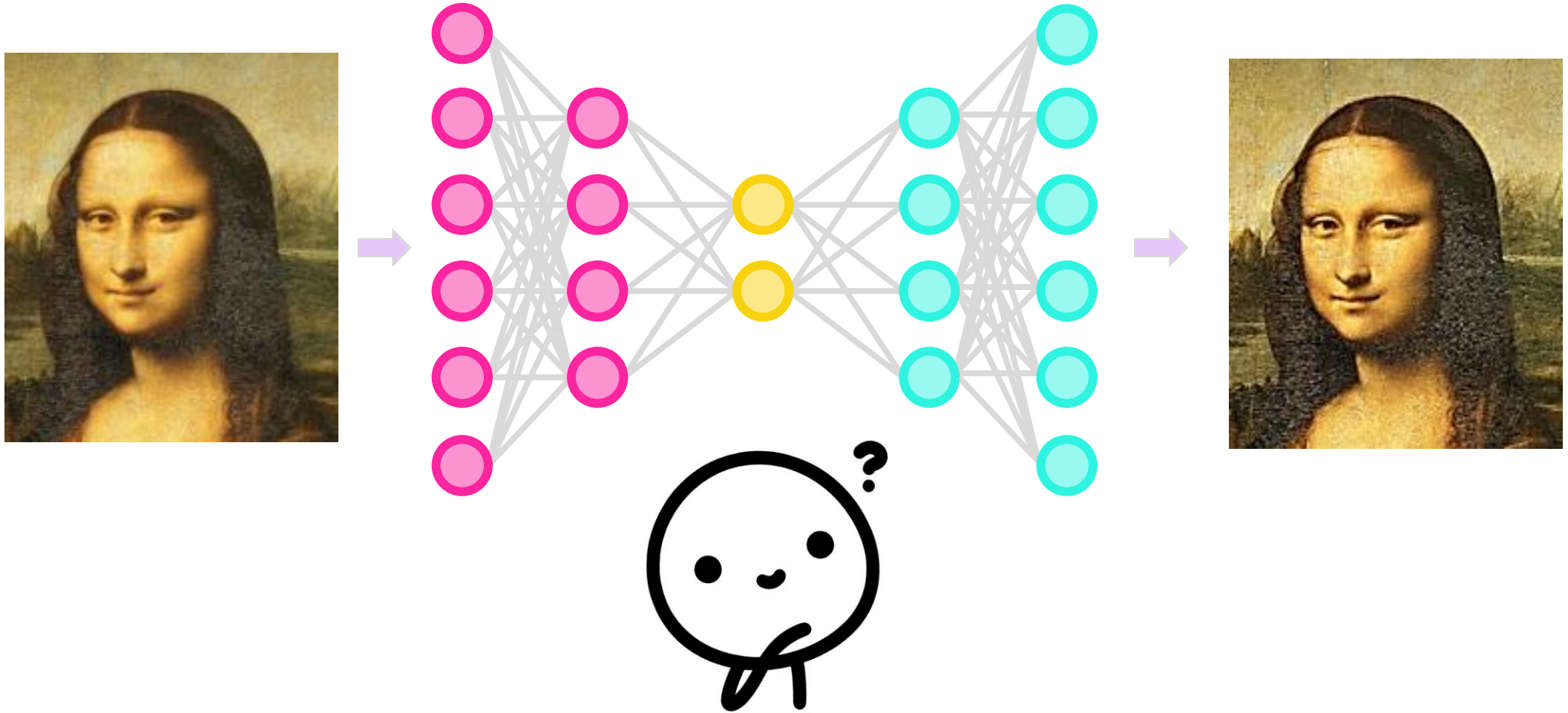
그렇게 함으로써 오토인코더는 주어진 데이터의 중요한 특성을 잡아내게 됩니다.

$$\left(\text{Original Image} - \text{Reconstructed Image} \right) = \text{Reconstruction Loss}$$

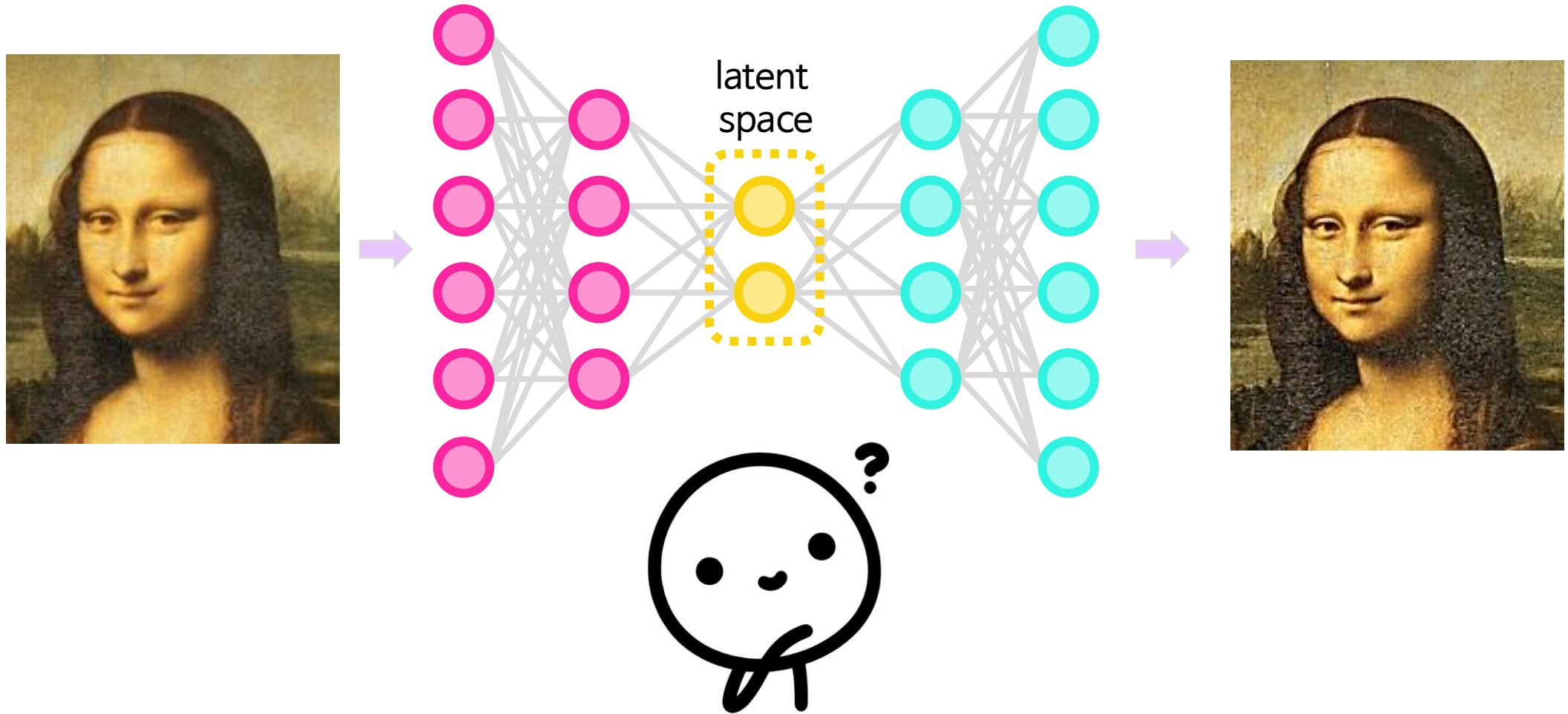
그렇다면, 오토인코더는 결국 입력을 받아서 그 입력과 거의 똑 같은
출력을 만들어내는 신경망인데요



결국 똑 같은 것을 만들어 낼 뿐인데, 도대체 어떻게 데이터의 중요한 특성을 잡아내게 되는 걸까요?



그 비밀은 잠재공간에서 일어나는 차원 축소 dimension reduction
때문입니다.



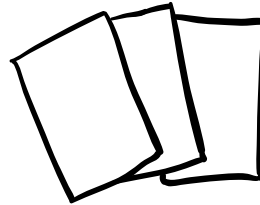
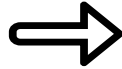
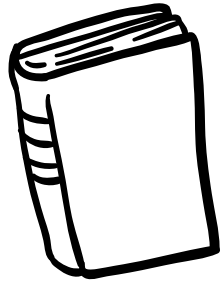
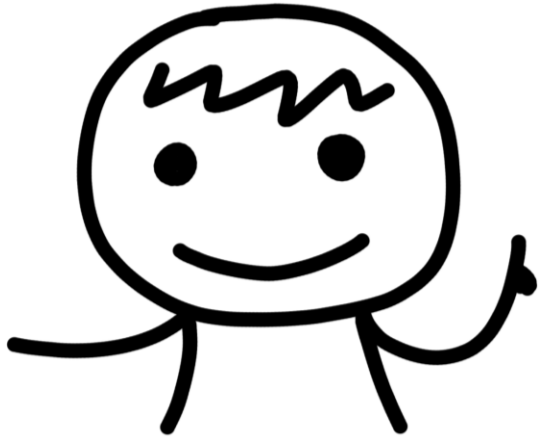
그 비밀은 잠재공간에서 일어나는 차원 축소 dimension reduction
때문입니다.

차원 축소란, 데이터의 특성을 줄이는 과정입니다. 이는 계산 효율성을 높이고,
노이즈를 줄이며, 데이터를 더 잘 이해하기 위해 사용됩니다. 예를 들어, 불필요한
정보를 제거하고 중요한 정보를 유지함으로써 모델의 성능을 개선할 수 있습니다.



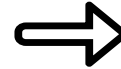
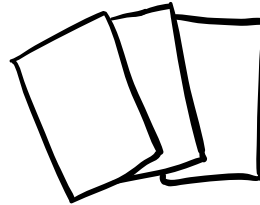
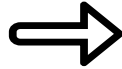
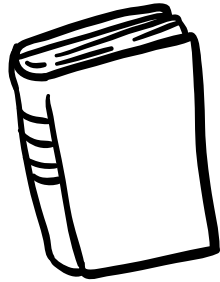
그 비밀은 잠재공간에서 일어나는 차원 축소 dimension reduction
때문입니다.

예를 들어, 200페이지의 책을 3페이지 요약본으로 만들어야 한다고 가정해봅시다.
원래의 책에는 많은 페이지와 정보가 있지만, 요약본은 핵심 내용만 간추려야 합니다.
이런 과정을 통해서 중요한 특징만 남게 됩니다.



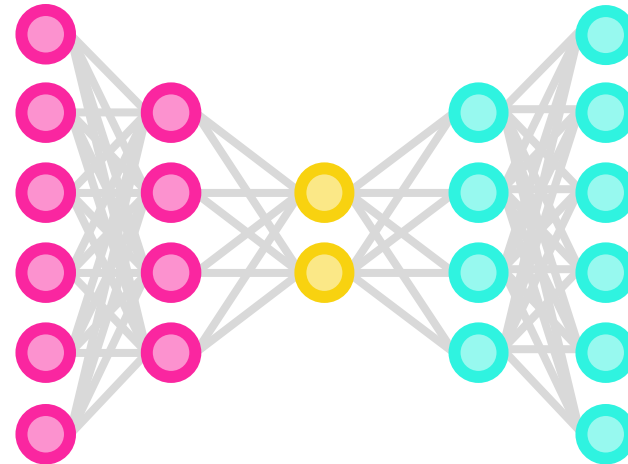
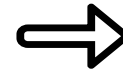
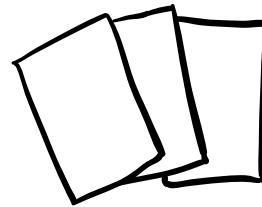
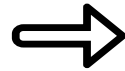
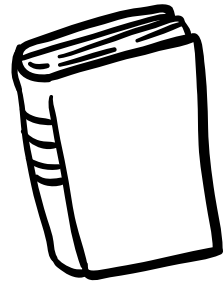
그 비밀은 잠재공간에서 일어나는 차원 축소 dimension reduction
때문입니다.

뿐만 아니라, 중요한 특징만 남은 요약본을 통해서 새로운 형태의 원본과 비슷한 책을
'생성'할 수도 있습니다.



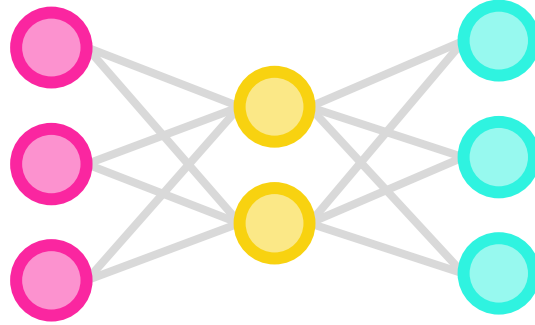
그 비밀은 잠재공간에서 일어나는 차원 축소 dimension reduction
때문입니다.

이것이 오토인코더가 하는 일을 비유적으로 보여주는 사례입니다. 이처럼 오토인코더는
원본 데이터를 제한된 크기의 잠재 공간을 사용하여 최대한 근접하게 생성해내는
학습을 하는 과정에서 원본데이터가 갖는 중요한 특성을 잡아내게 됩니다

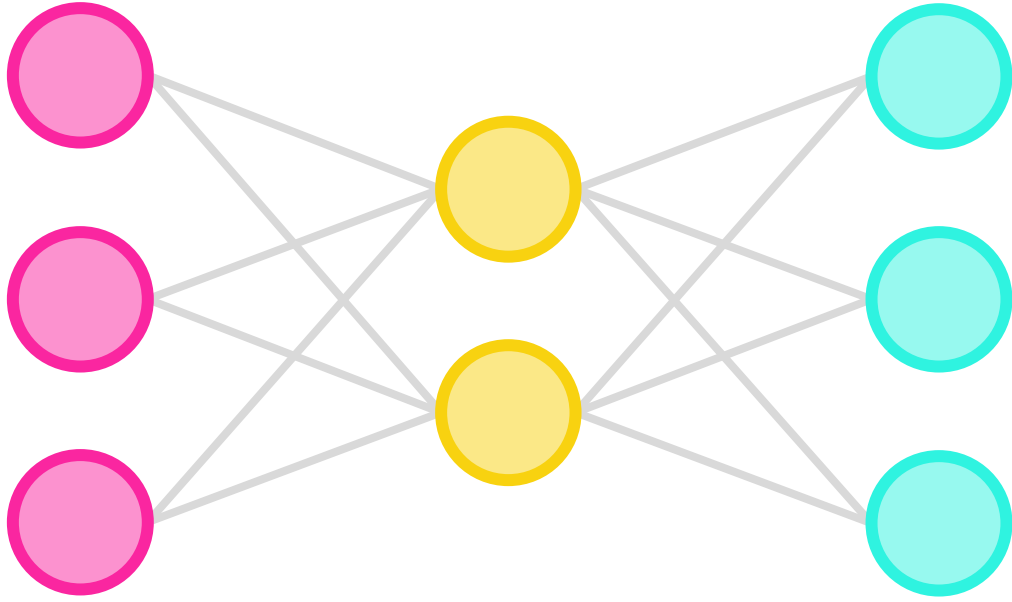


자 그러면 이제 예제를 통해 오토인코더의 학습에 대해 알아보도록 하겠습니다.

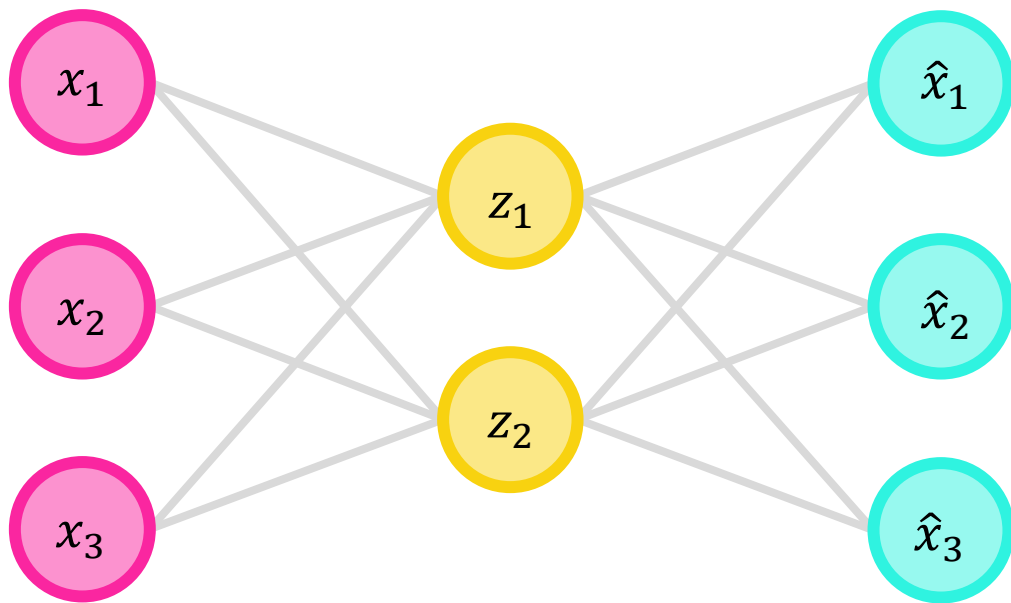
보다 쉬운 계산을 위해 간단한 모델을 가정해보겠습니다.



보다 쉬운 계산을 위해 간단한 모델을 가정해보겠습니다.



먼저, 다음과 같이 초기값을 설정하겠습니다.

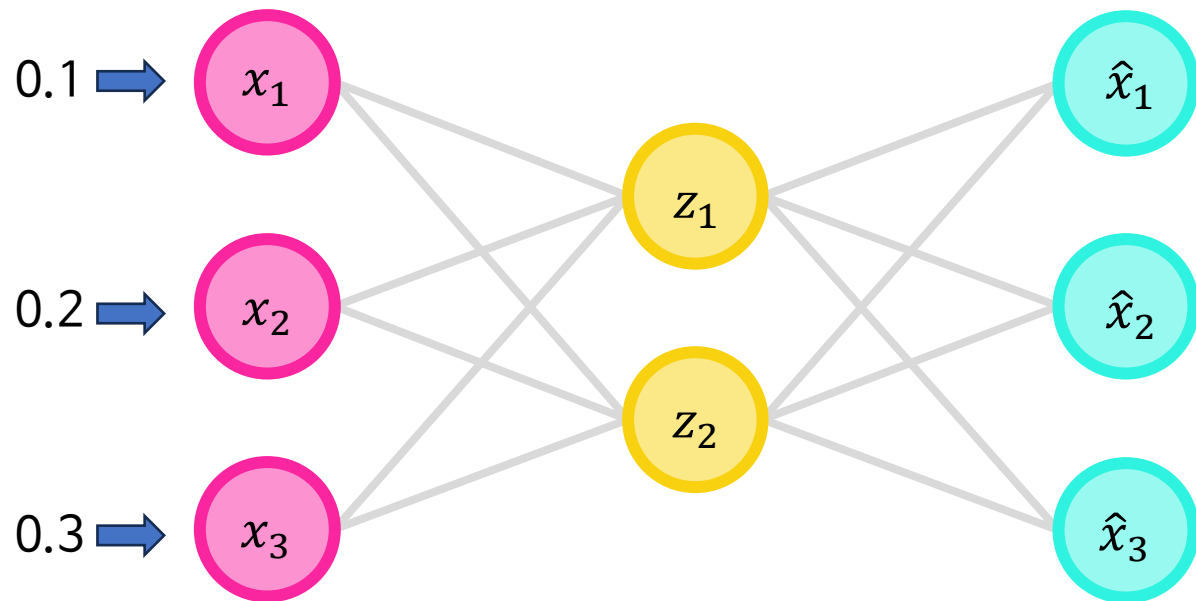


입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

다음과 같이 입력값을 넣을 경우,

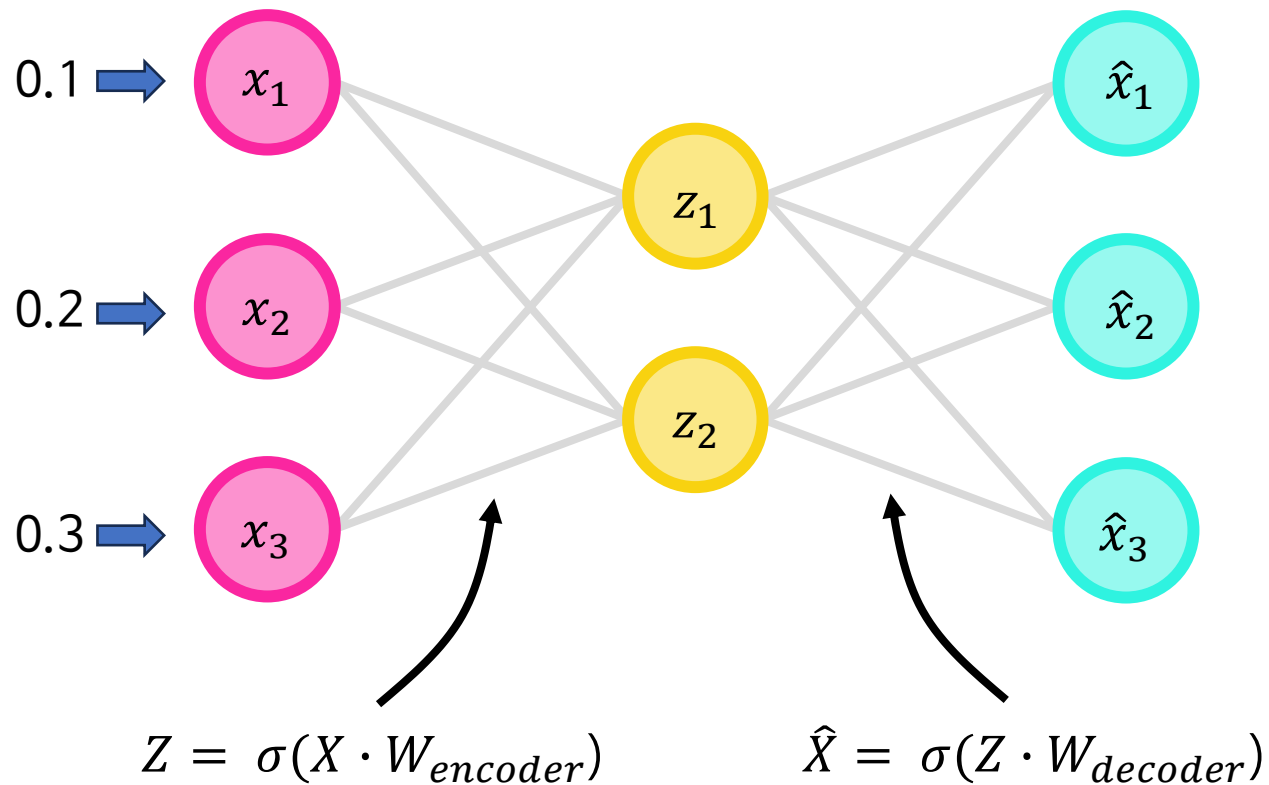


입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

순전파 feedforward 계산은 다음과 같이 두식을 차례로 계산합니다

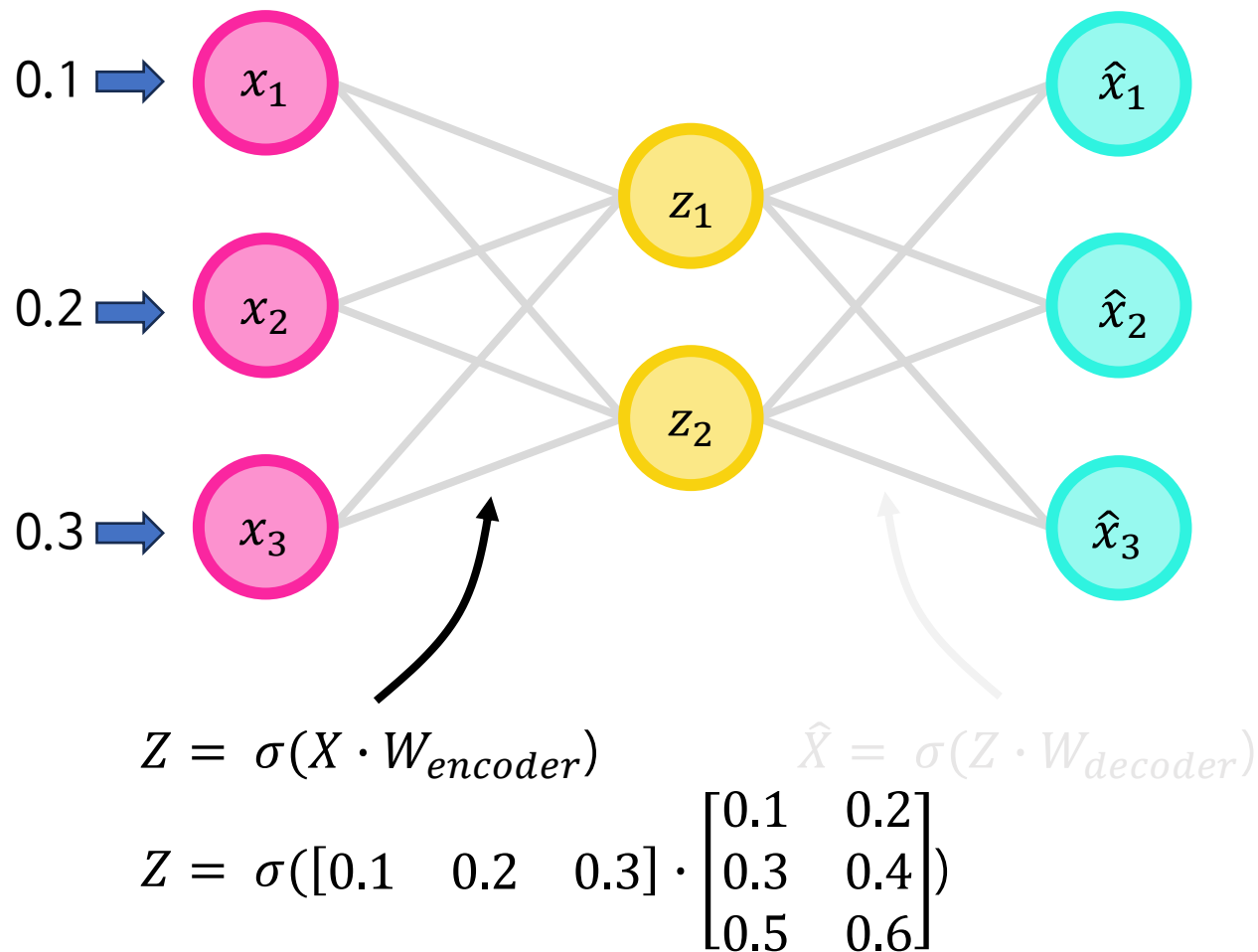


입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

먼저 이 부분을 계산하면,

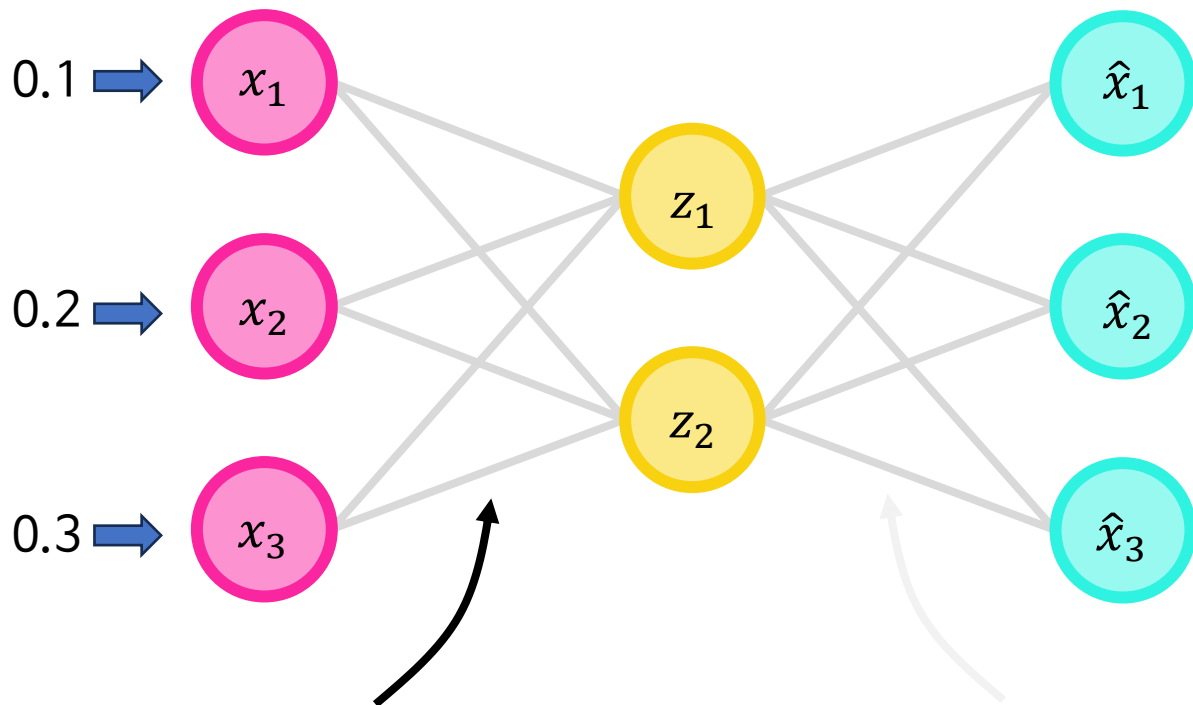


입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

이렇게 계산이 되며,



$$Z = \sigma(X \cdot W_{encoder})$$

$$Z = \sigma([0.1 \quad 0.2 \quad 0.3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([0.22 \quad 0.28])$$

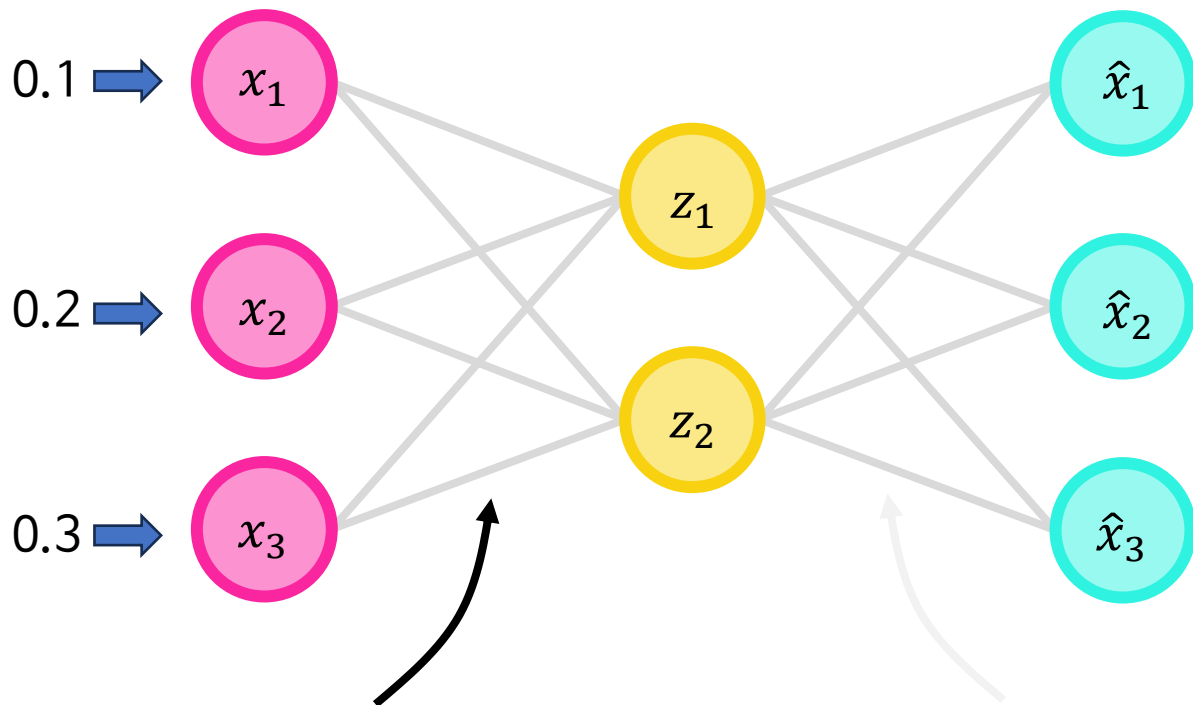
$$\hat{X} = \sigma(Z \cdot W_{decoder})$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

$$\text{디코더가중치: } W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

이렇게 z값을 구할 수가 있고,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

$$Z = \sigma(X \cdot W_{encoder})$$

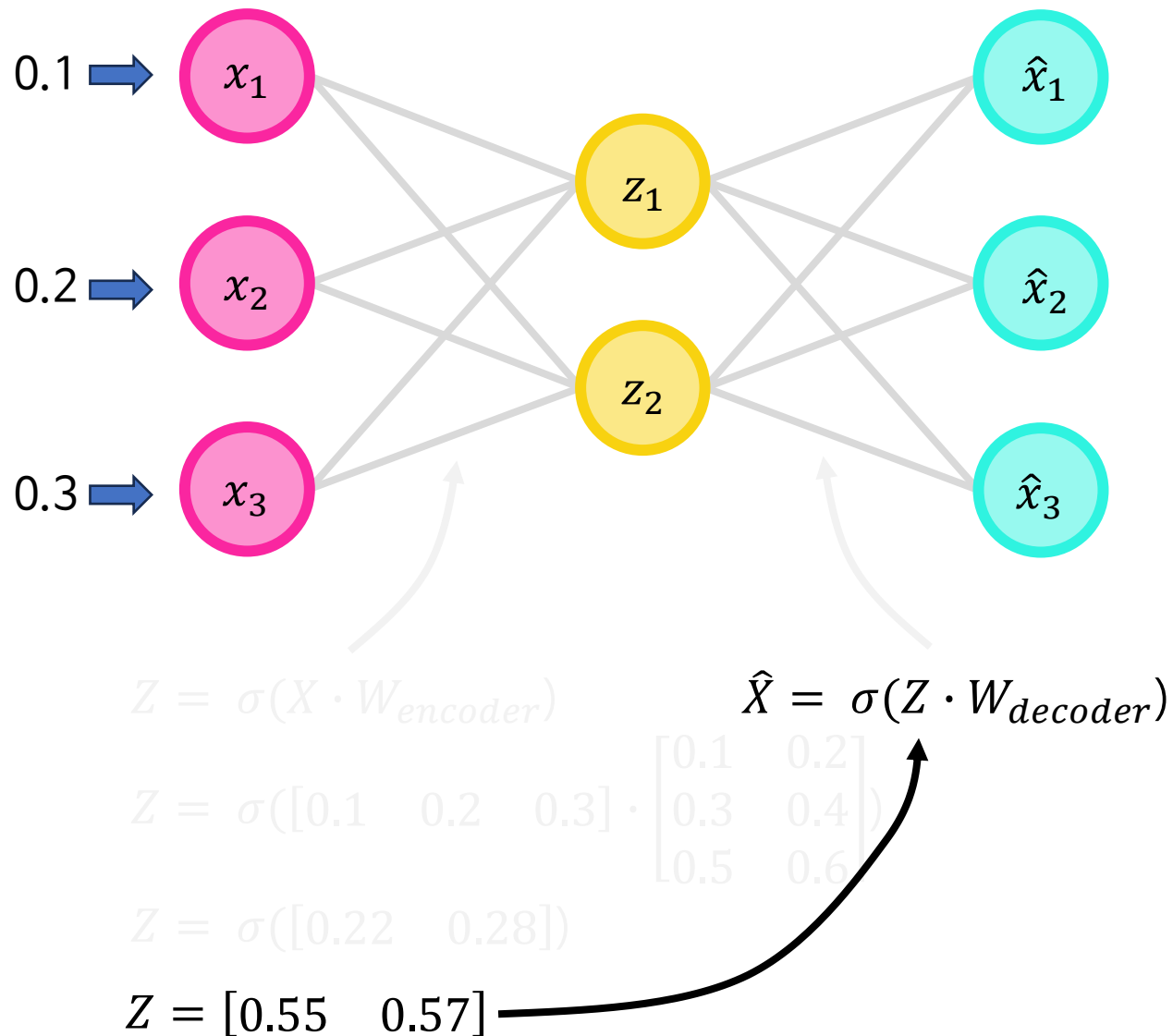
$$Z = \sigma([0.1 \quad 0.2 \quad 0.3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([0.22 \quad 0.28])$$

$$Z = [0.55 \quad 0.57]$$

$$\hat{X} = \sigma(Z \cdot W_{decoder})$$

계속해서 이 z값을 여기에 대입합니다



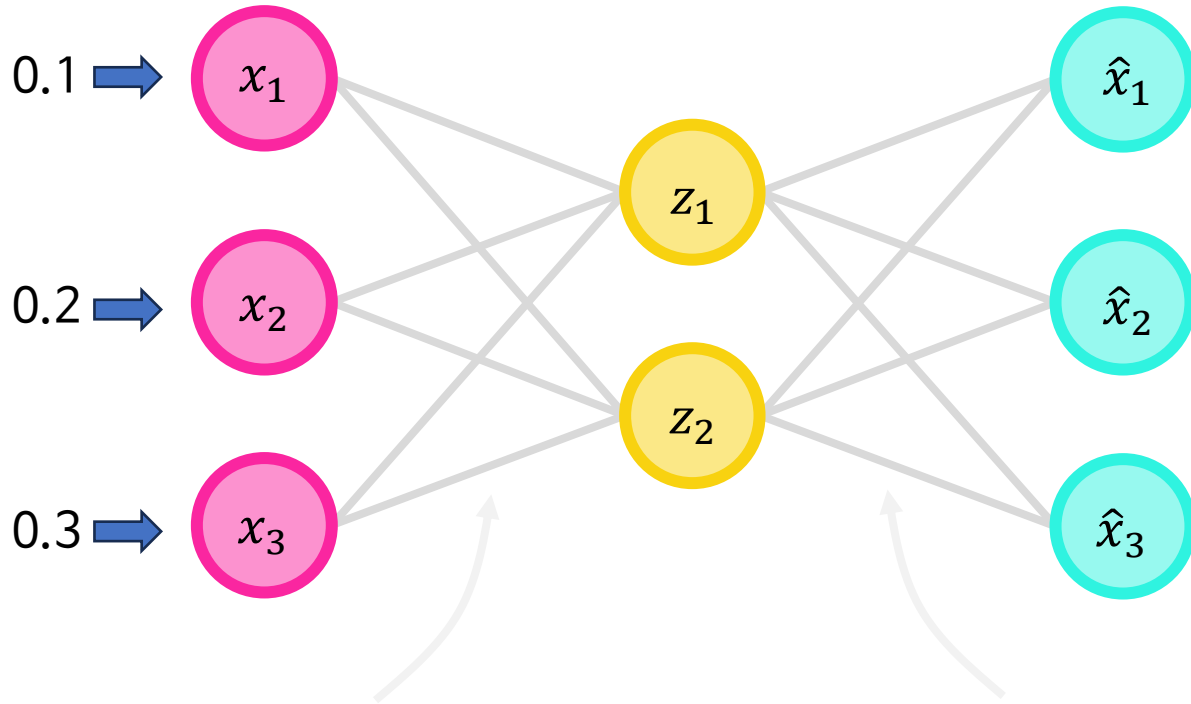
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

이렇게 디코더 가중치를 대입하고,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$



$$\hat{X} = \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix})$$

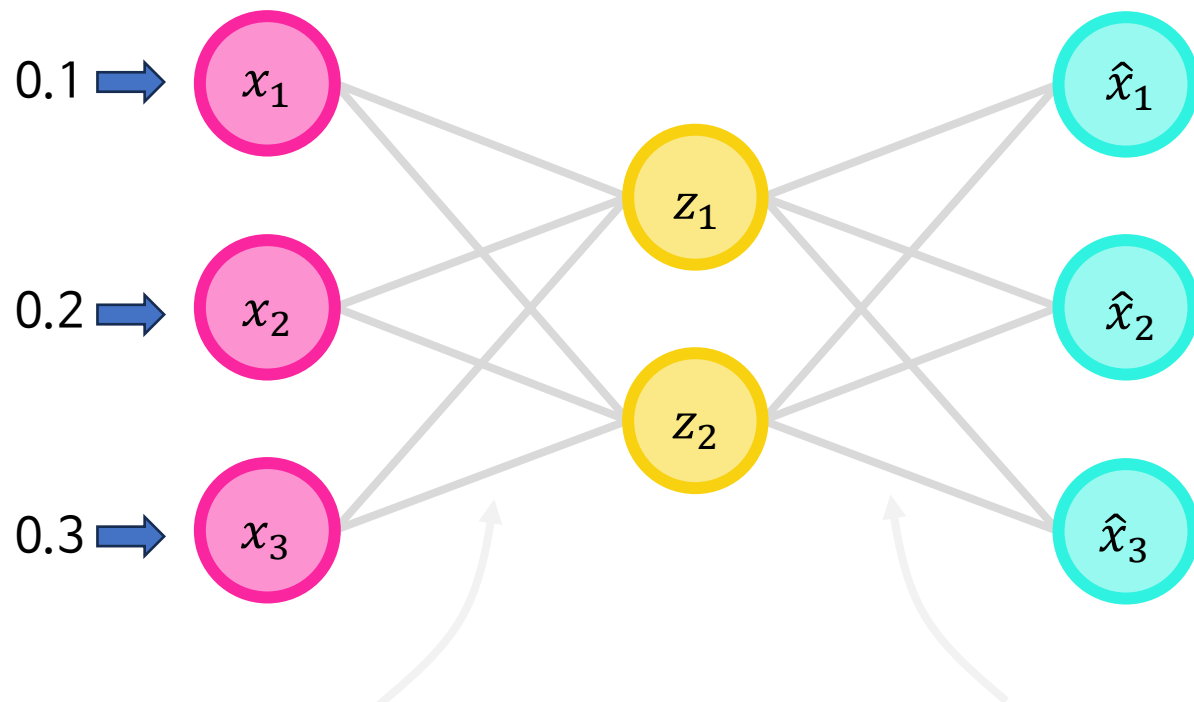
$$Z = \sigma(X \cdot W_{encoder})$$

$$Z = \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([2.2 \quad 2.8])$$

$$Z = [0.9 \quad 0.94]$$

계속해서 계산을 하면,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$Z = \sigma(X \cdot W_{encoder})$$

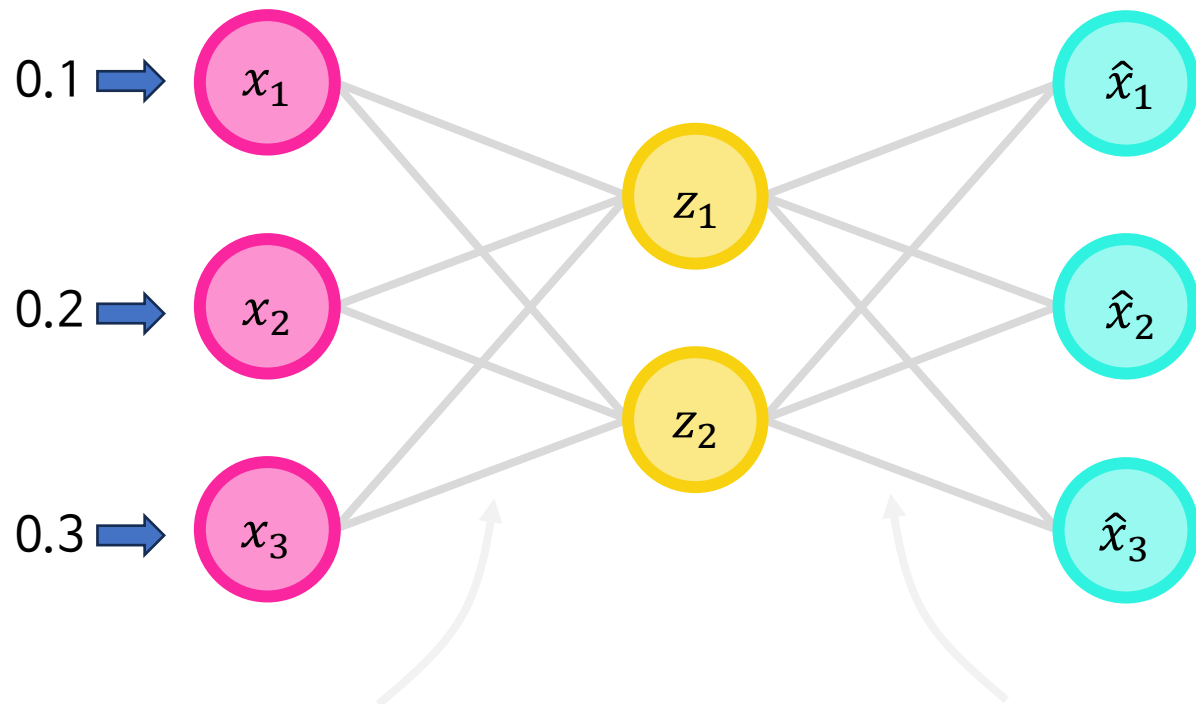
$$Z = \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([2.2 \quad 2.8])$$

$$Z = [0.9 \quad 0.94]$$

$$\begin{aligned} \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\ &= \sigma([0.28 \quad 0.39 \quad 0.51]) \end{aligned}$$

\hat{X} 값은 이렇게 계산할 수 있습니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$Z = \sigma(X \cdot W_{encoder})$$

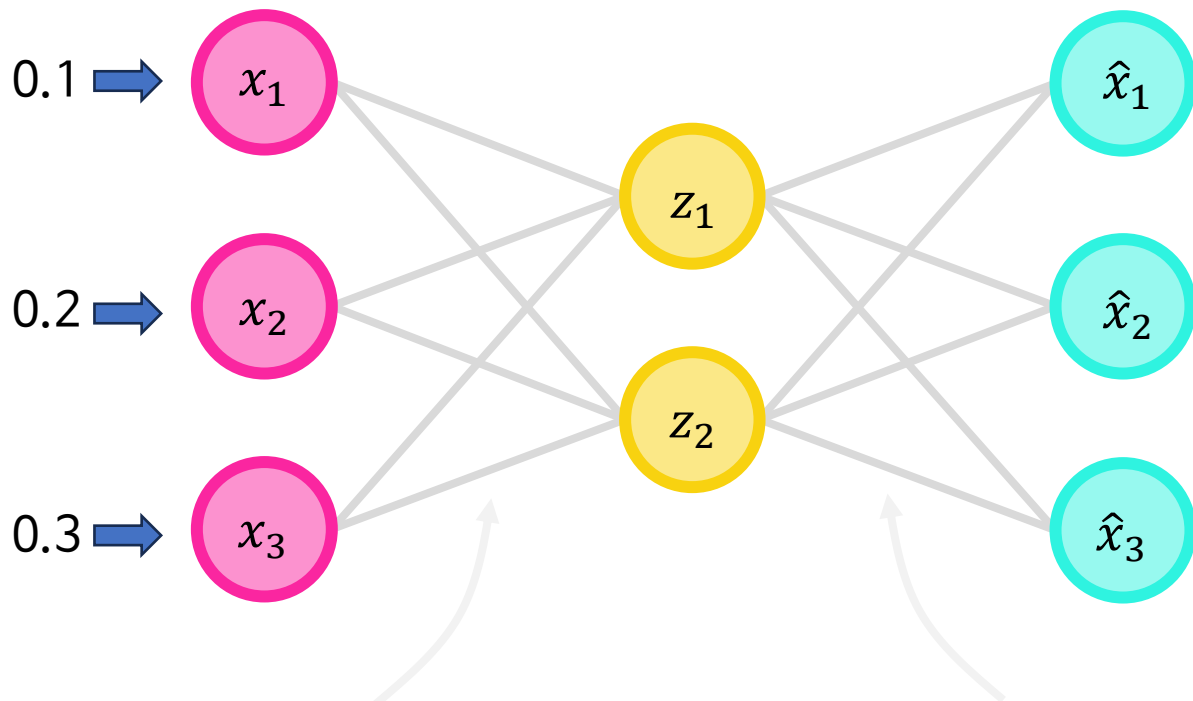
$$Z = \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([2.2 \quad 2.8])$$

$$Z = [0.9 \quad 0.94]$$

$$\begin{aligned} \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\ &= \sigma([0.28 \quad 0.39 \quad 0.51]) \\ &= [0.57 \quad 0.60 \quad 0.62] \end{aligned}$$

이렇게 하면 순전파 feedforward는 계산 완료입니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$Z = \sigma(X \cdot W_{encoder})$$

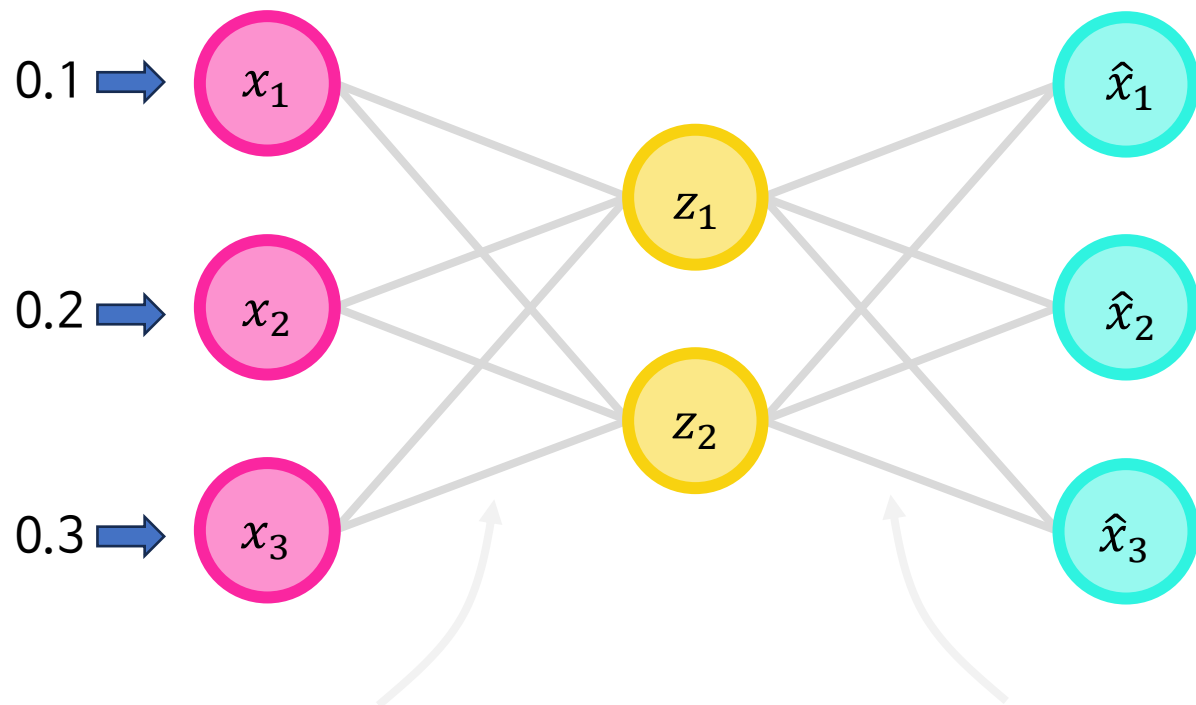
$$Z = \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([2.2 \quad 2.8])$$

$$Z = [0.9 \quad 0.94]$$

$$\begin{aligned} \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\ &= \sigma([0.28 \quad 0.39 \quad 0.51]) \\ &= [0.57 \quad 0.60 \quad 0.62] \end{aligned}$$

그러면 이제 손실loss를 구해볼 차례입니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$Z = \sigma(X \cdot W_{encoder})$$

$$Z = \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix})$$

$$Z = \sigma([2.2 \quad 2.8])$$

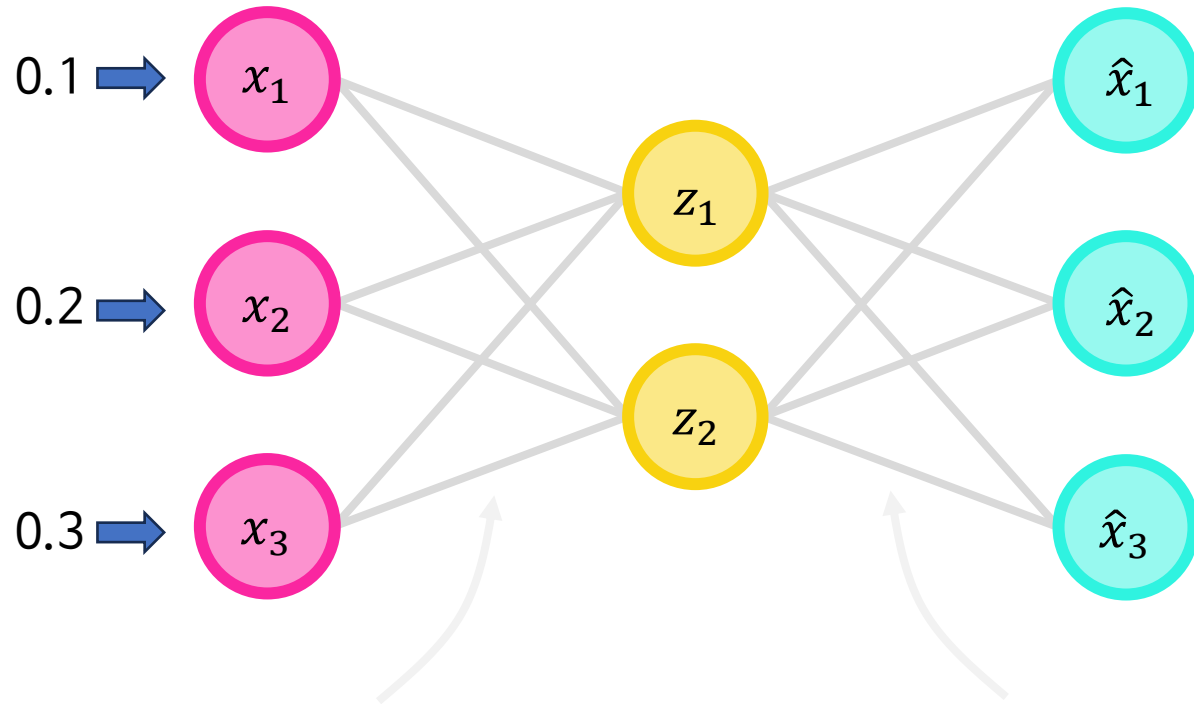
$$Z = [0.9 \quad 0.94]$$

$$\hat{X} = \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix})$$

$$= \sigma([0.28 \quad 0.39 \quad 0.51])$$

$$= [0.57 \quad 0.60 \quad 0.62]$$

오토인코더의 손실 계산법은 너무너무 간단합니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

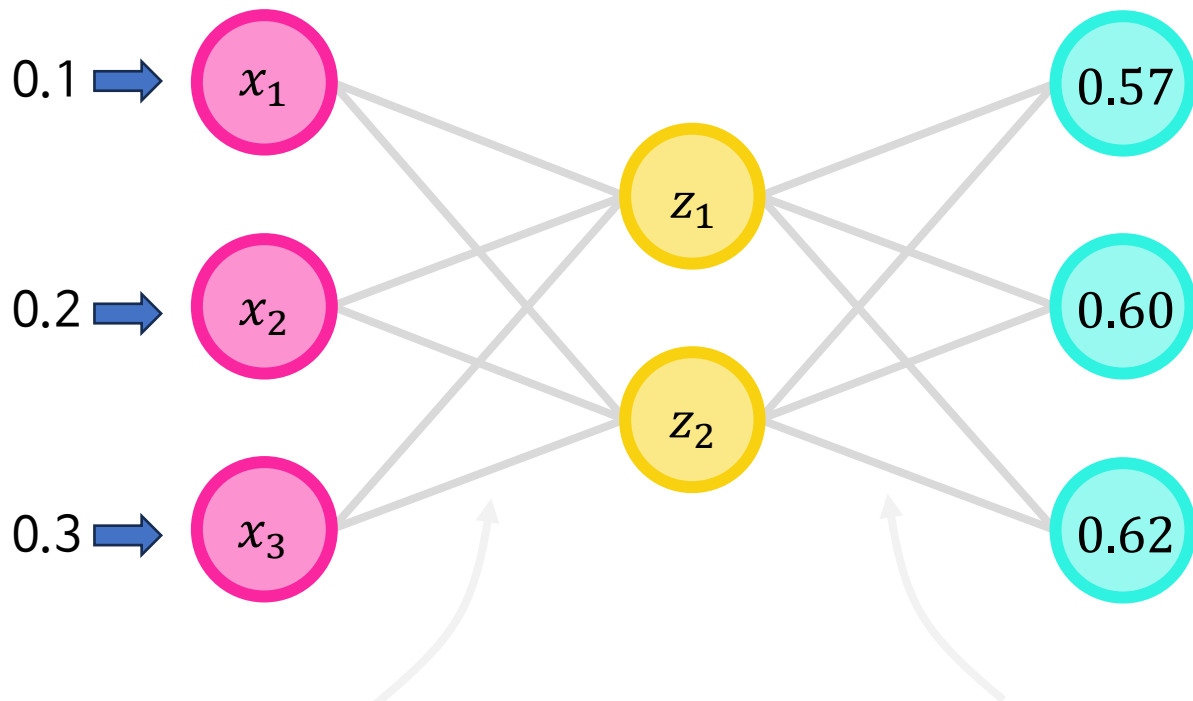
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$\begin{aligned} Z &= \sigma(X \cdot W_{encoder}) & \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\ Z &= \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}) & &= \sigma([0.28 \quad 0.39 \quad 0.51]) \\ Z &= \sigma([2.2 \quad 2.8]) & &= [0.57 \quad 0.60 \quad 0.62] \\ Z &= [0.9 \quad 0.94] \end{aligned}$$

이렇게 출력값을 대입하고



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

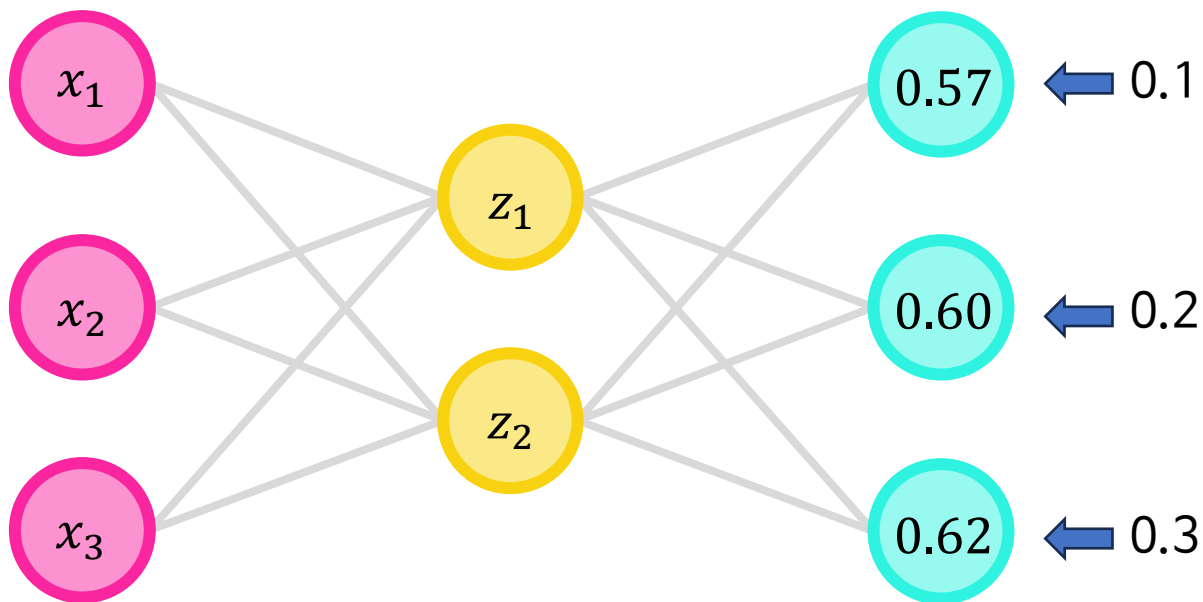
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$\begin{aligned}
 Z &= \sigma(X \cdot W_{encoder}) \\
 Z &= \sigma([1 \quad 2 \quad 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}) \\
 Z &= \sigma([2.2 \quad 2.8]) \\
 Z &= [0.9 \quad 0.94] \\
 \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.55 \quad 0.57] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\
 &= \sigma([0.28 \quad 0.39 \quad 0.51]) \\
 &= [0.57 \quad 0.60 \quad 0.62]
 \end{aligned}$$

이렇게 입력값을 배치하면,



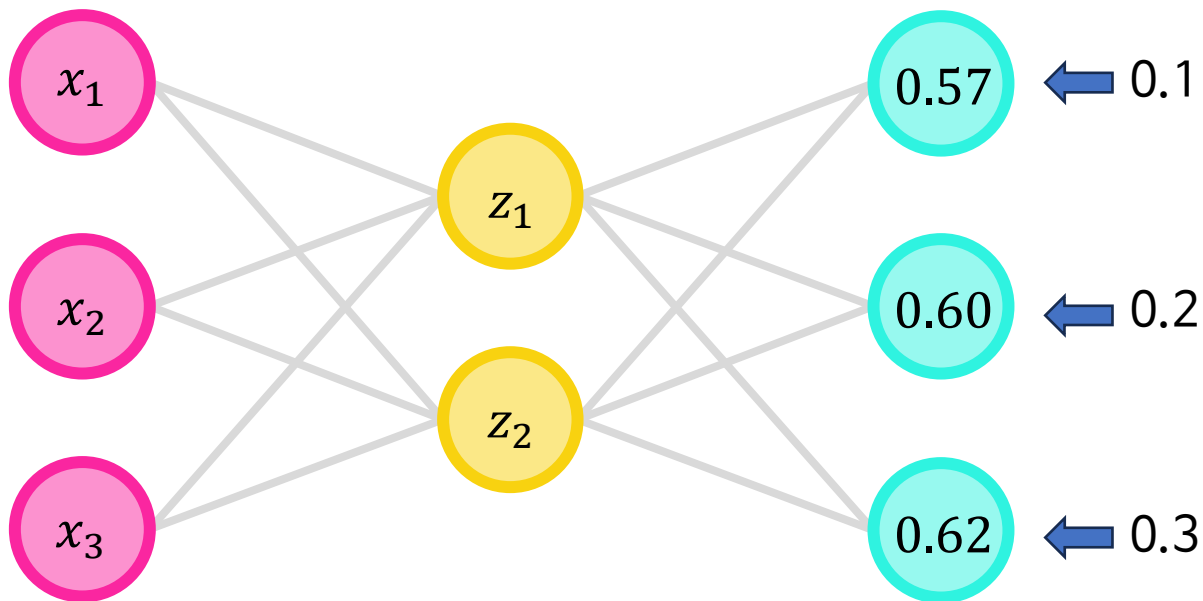
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

이렇게 입력값과 출력값을 비교해보도록 하겠습니다



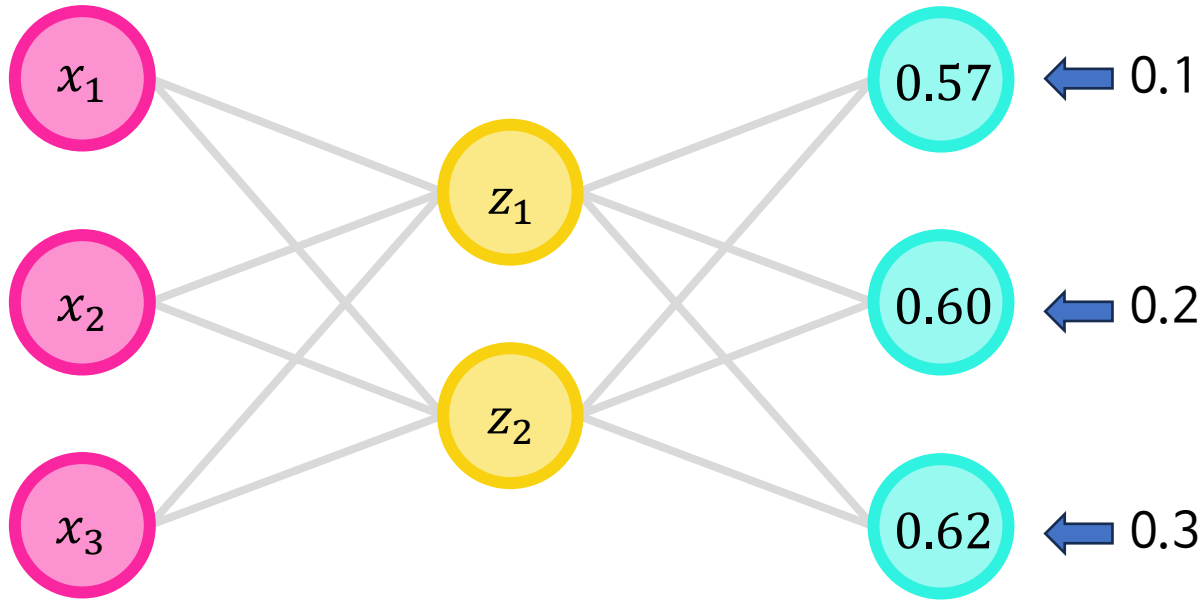
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

손실함수는 간단한 MSE를 사용하도록 하겠습니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

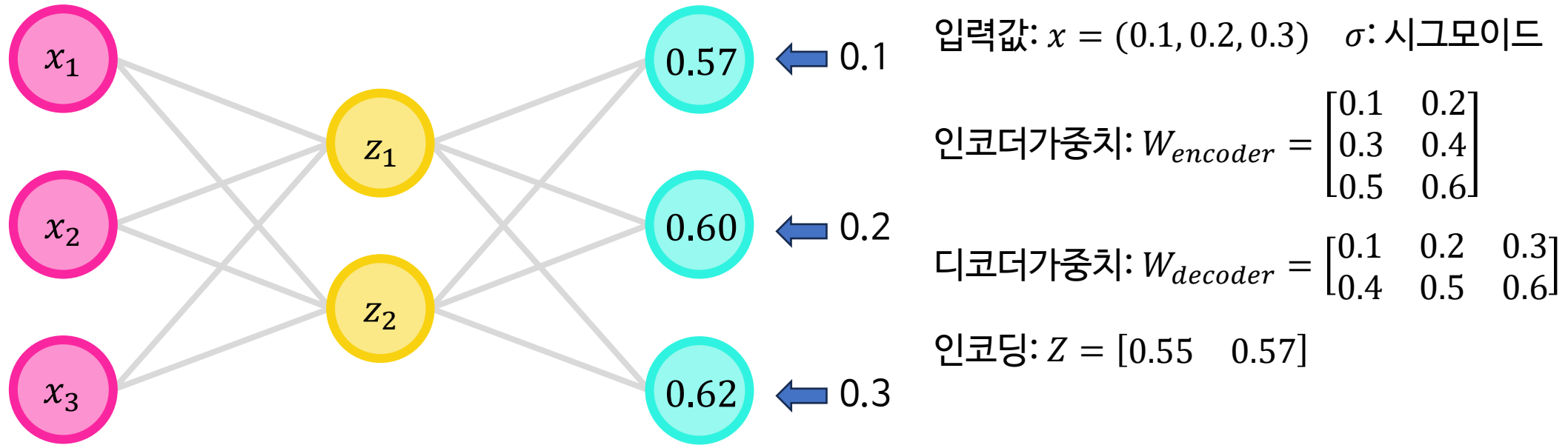
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

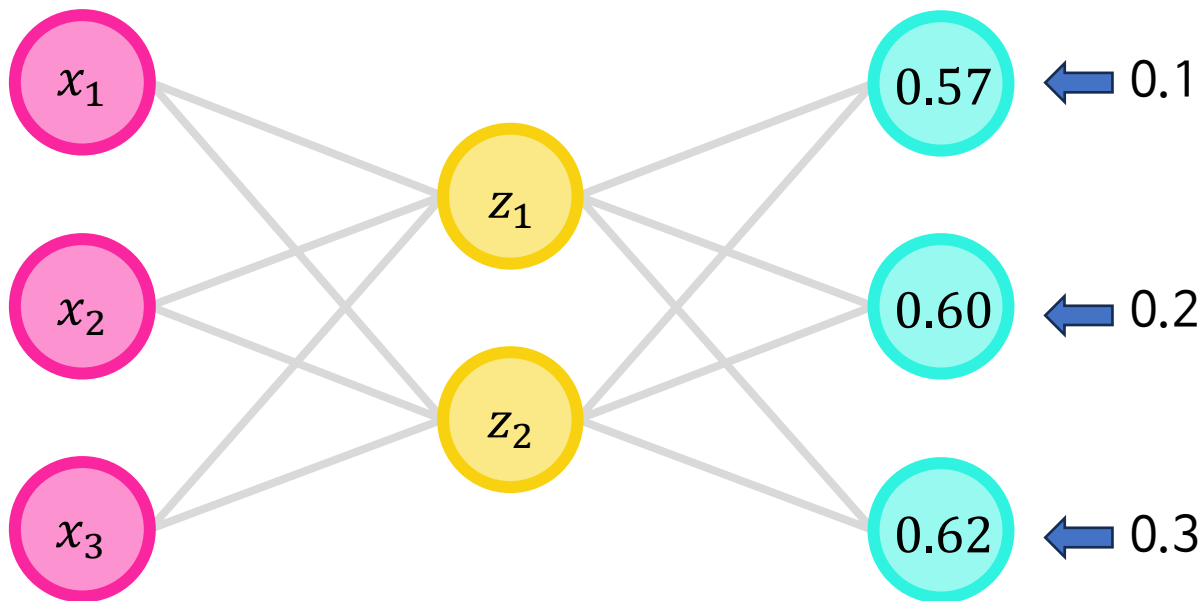
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

손실함수는 간단한 MSE를 사용하도록 하겠습니다.



$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{3} \{ (0.1 - 0.57)^2 + (0.2 - 0.60)^2 + (0.3 - 0.62)^2 \}$$

손실 loss는 0.16가 나왔습니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

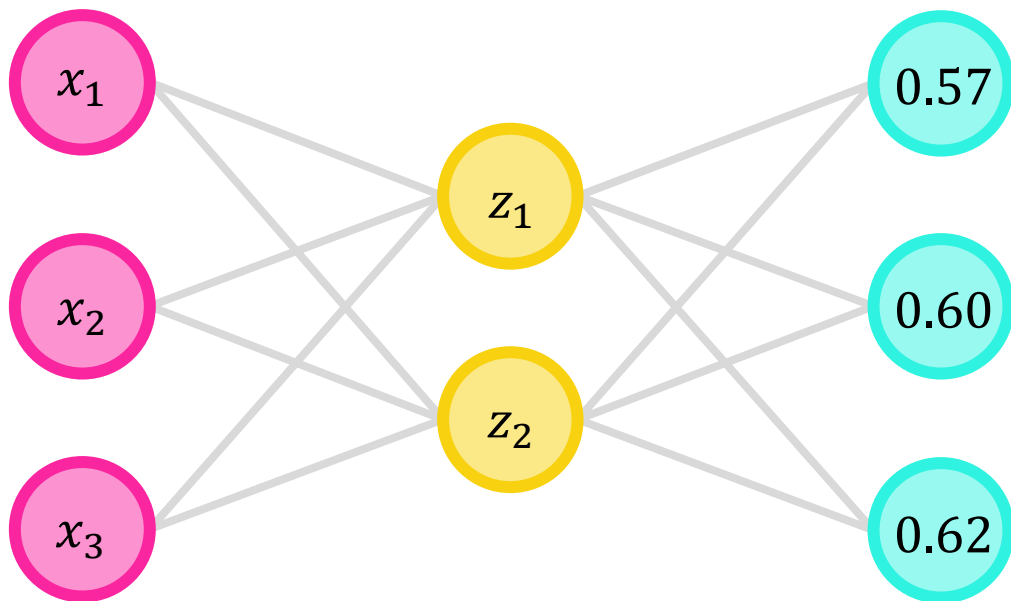
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{3} \{ (0.1 - 0.57)^2 + (0.2 - 0.60)^2 + (0.3 - 0.62)^2 \} = 0.16$$

자 이제는 역전파를 이용해 각각의 가중치를 업데이트 해보겠습니다



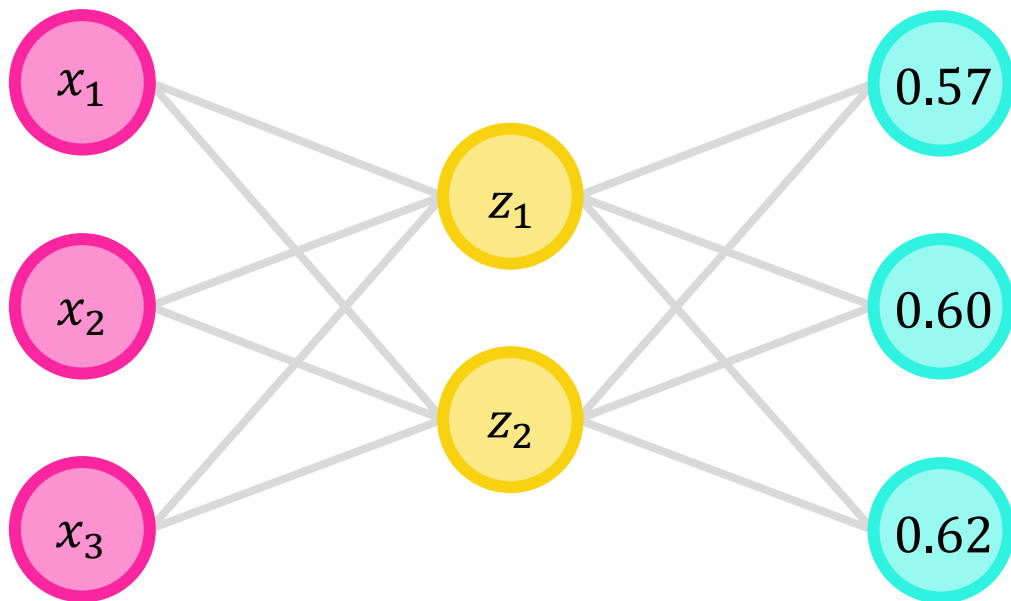
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

먼저 $W_{decoder}$ 의 gradient를 구해보도록 하겠습니다



$$\frac{\partial L}{\partial W_{decoder}}$$

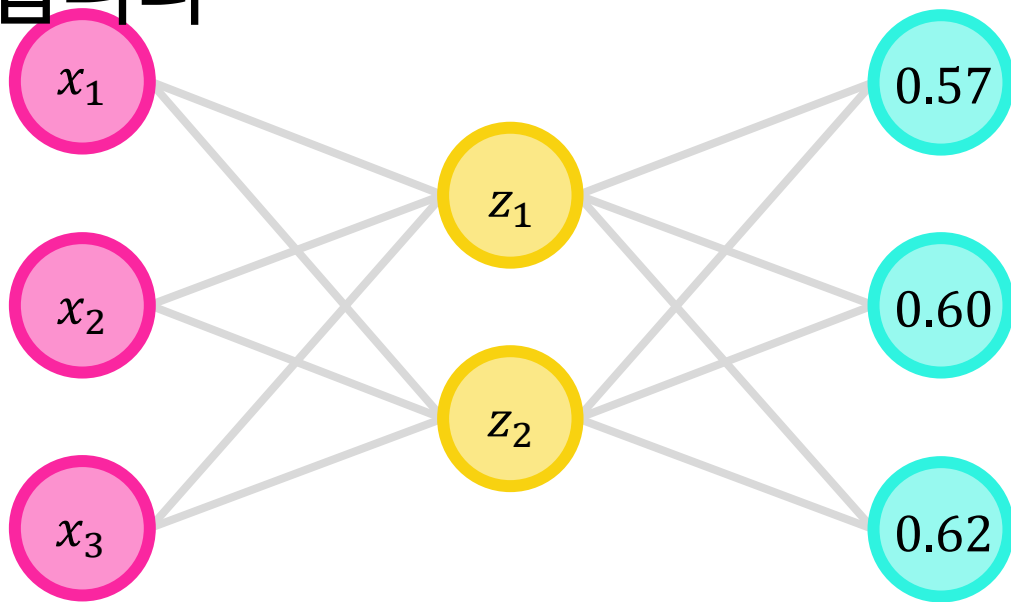
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

왜냐하면, gradient는 경사하강법의 가중치 업데이트에 필요하기 때문입니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

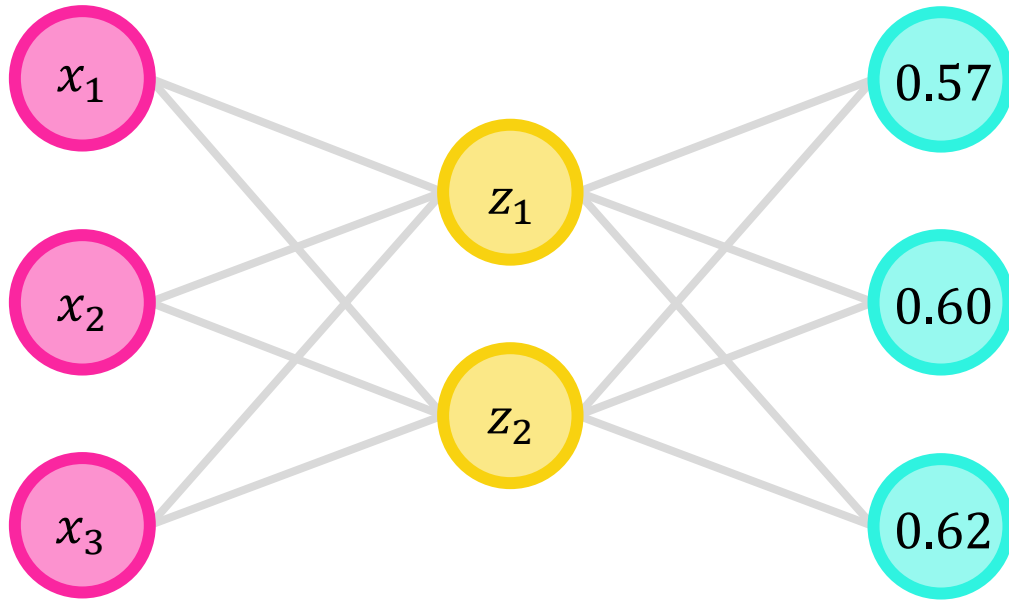
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$W_{decoder}^* = W_{decoder} - \alpha \cdot \frac{\partial L}{\partial W_{decoder}}$$

$\partial L / \partial W_{decoder}$ 는 다음과 같이 전개할 수 있습니다.



$$\frac{\partial L}{\partial W_{decoder}}$$

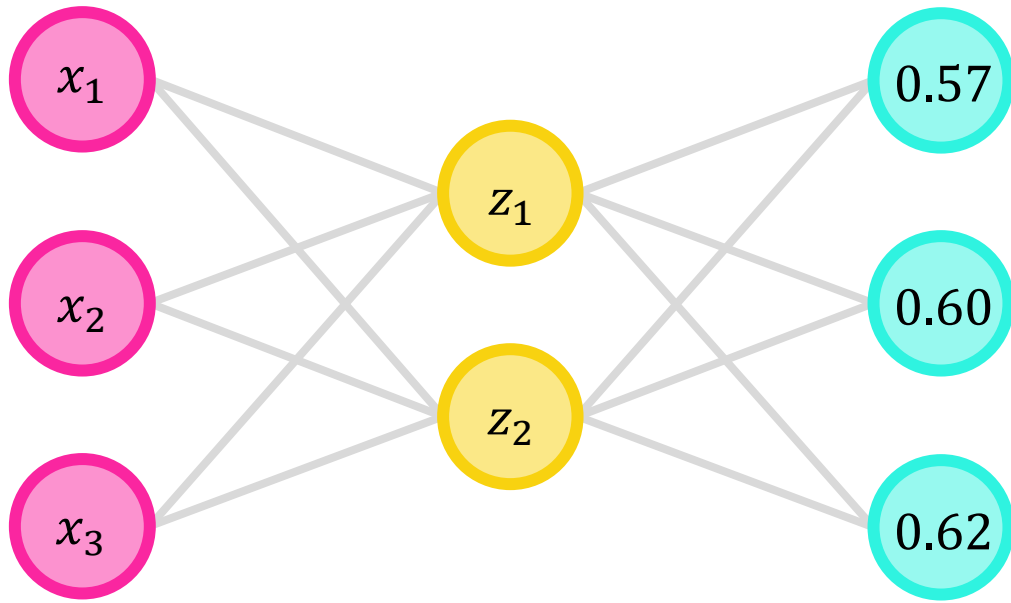
입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$\partial L / \partial W_{decoder}$ 는 다음과 같이 전개할 수 있습니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

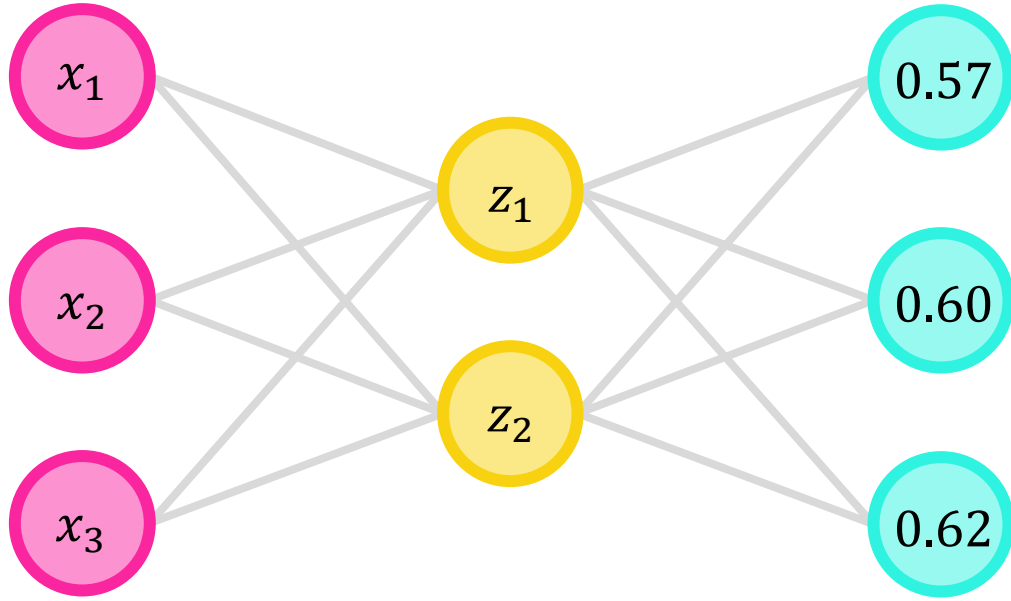
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}}$$

그리고 MSE의 공식에 의해서 $\partial L / \partial \hat{x}$ 는 다음과 같이 바뀌고,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

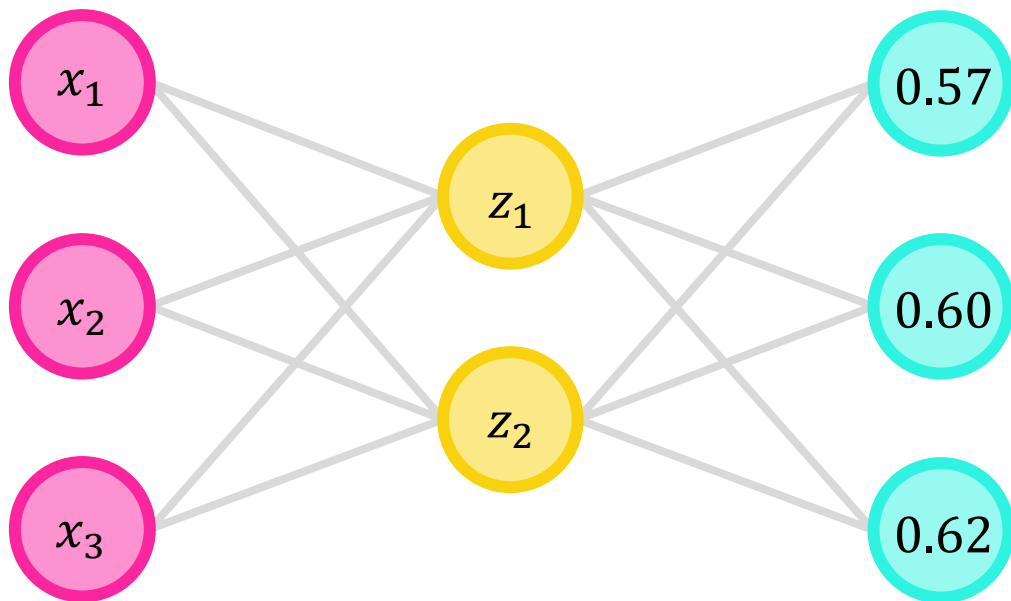
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial W_{decoder}}$$

다음 \hat{x} 공식에 의해서 $\partial \hat{x} / \partial W_{decoder}$ 은 다음과 같이 바뀝니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

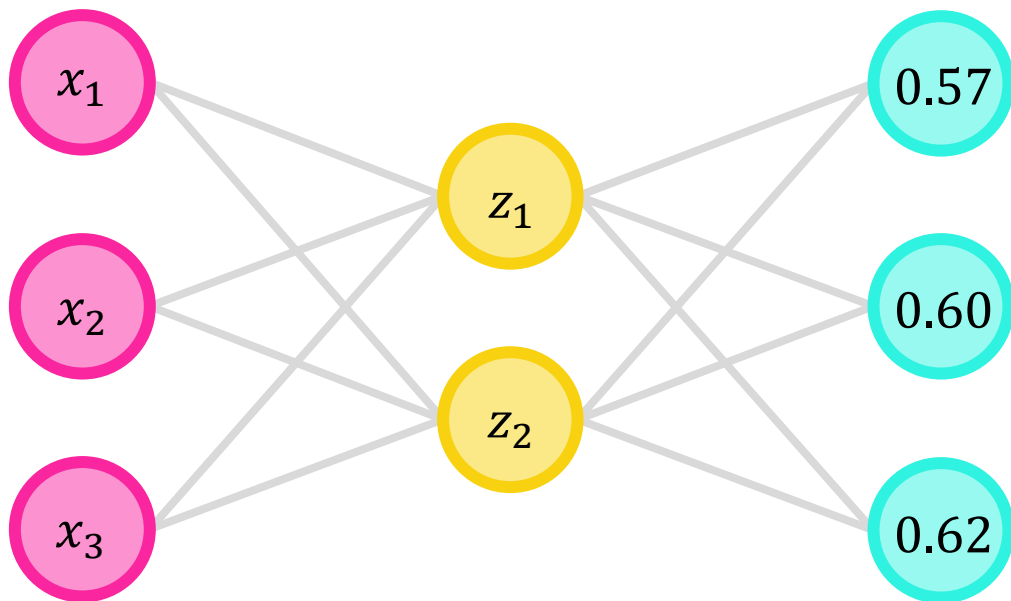
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\hat{x} = \sigma(z \cdot W_{decoder})$$

$$\frac{\partial L}{\partial W_{decoder}} = \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial W_{decoder}}$$

다음 \hat{x} 공식에 의해서 $\partial \hat{x} / \partial W_{decoder}$ 은 다음과 같이 바뀝니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

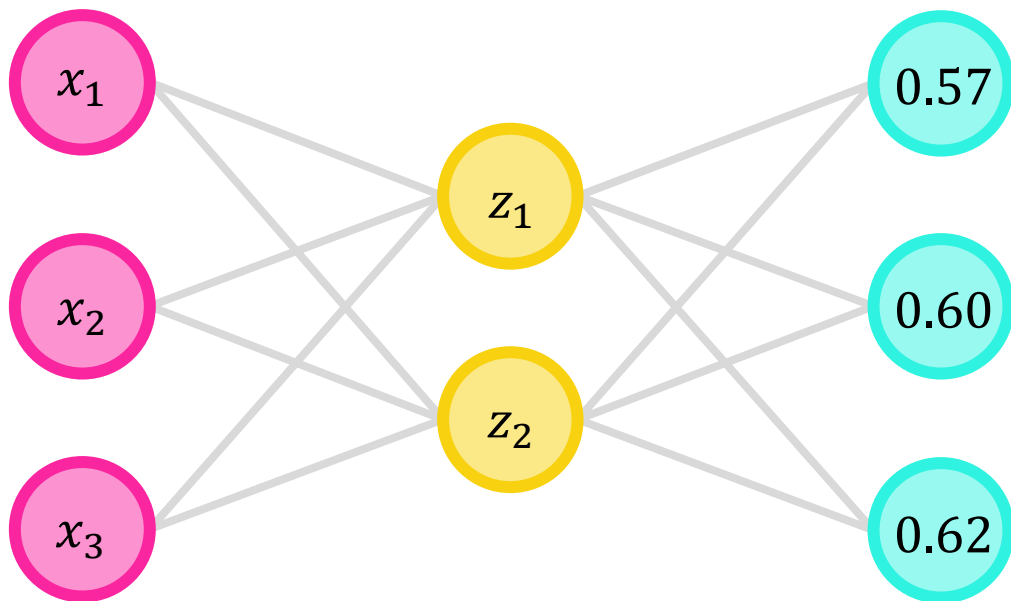
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\hat{x} = \sigma(z \cdot W_{decoder})$$

$$\frac{\partial L}{\partial W_{decoder}} = \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \sigma(z \cdot W_{decoder})}{\partial W_{decoder}}$$

다음 \hat{x} 공식에 의해서 $\partial \hat{x} / \partial W_{decoder}$ 은 다음과 같이 바뀝니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

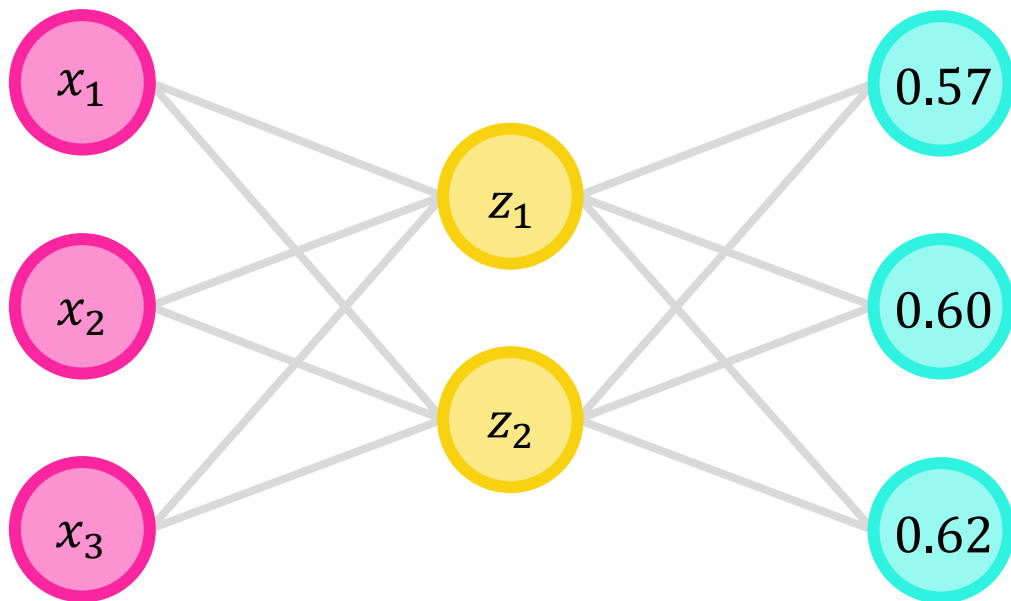
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\begin{aligned}
 \frac{\partial L}{\partial W_{decoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \sigma(z \cdot W_{decoder})}{\partial W_{decoder}} \\
 &= 2 \cdot (\hat{x} - x) \cdot z^T \cdot \sigma(z \cdot W_{decoder}) \cdot (1 - \sigma(z \cdot W_{decoder}))
 \end{aligned}$$

$\hat{x} = \sigma(z \cdot W_{decoder})$
 $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$

다음 \hat{x} 공식에 의해서 $\partial \hat{x} / \partial W_{decoder}$ 은 다음과 같이 바뀝니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

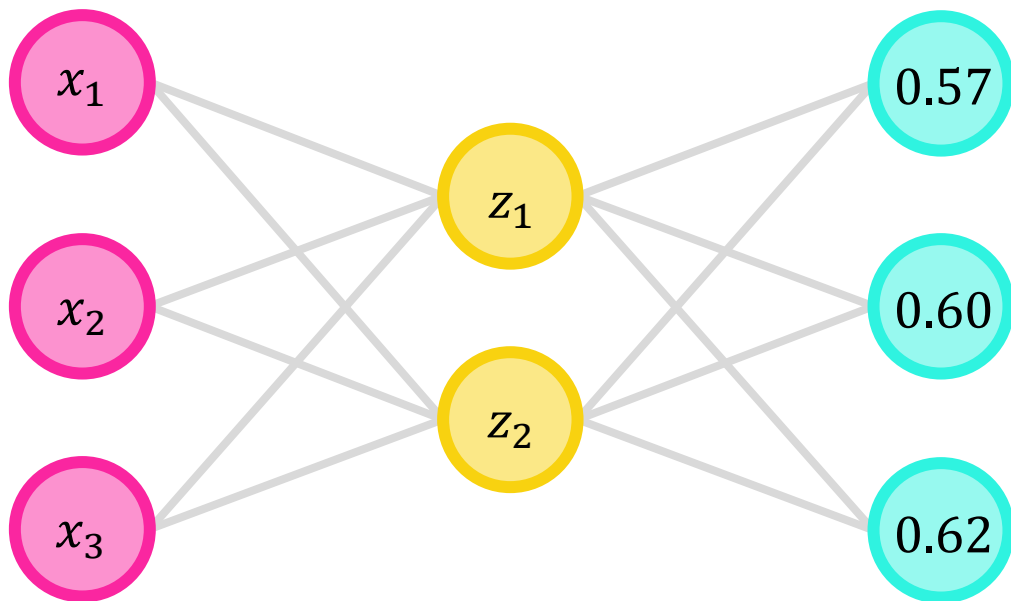
인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\begin{aligned}
 \frac{\partial L}{\partial W_{decoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \sigma(z \cdot W_{decoder})}{\partial W_{decoder}} \\
 &= 2 \cdot (\hat{x} - x) \cdot z^T \cdot \sigma(z \cdot W_{decoder}) \cdot (1 - \sigma(z \cdot W_{decoder})) \\
 &= 2 \cdot (\hat{x} - x) \cdot z^T \cdot \hat{x} \cdot (1 - \hat{x})
 \end{aligned}$$

Annotations for the derivation:

- $\hat{x} = \sigma(z \cdot W_{decoder})$ (points to the $\sigma(z \cdot W_{decoder})$ term in the first line)
- $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$ (points to the $\sigma(z \cdot W_{decoder}) \cdot (1 - \sigma(z \cdot W_{decoder}))$ term in the second line)
- $\hat{x} = \sigma(z \cdot W_{decoder})$ (points to the \hat{x} term in the third line)

각각의 값을 넣어서 계산하면 다음과 같이 됩니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

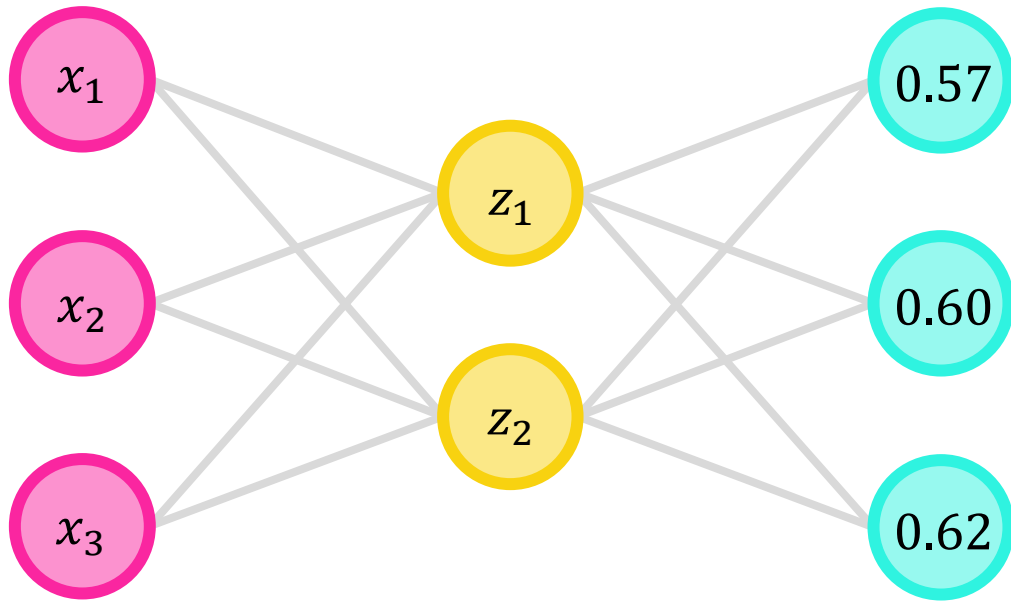
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\begin{aligned}
 \frac{\partial L}{\partial W_{decoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial W_{decoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \sigma(z \cdot W_{decoder})}{\partial W_{decoder}} \\
 &= 2 \cdot (\hat{x} - x) \cdot z^T \cdot \sigma(z \cdot W_{decoder}) \cdot (1 - \sigma(z \cdot W_{decoder})) \\
 &= 2 \cdot (\hat{x} - x) \cdot z^T \cdot \hat{x} \cdot (1 - \hat{x}) \\
 &= \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}
 \end{aligned}$$

$\hat{x} = \sigma(z \cdot W_{decoder})$
 $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$

마찬가지로 $W_{encoder}$ 의 기울기를 구해보도록 하겠습니다



$$\frac{\partial L}{\partial W_{encoder}}$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

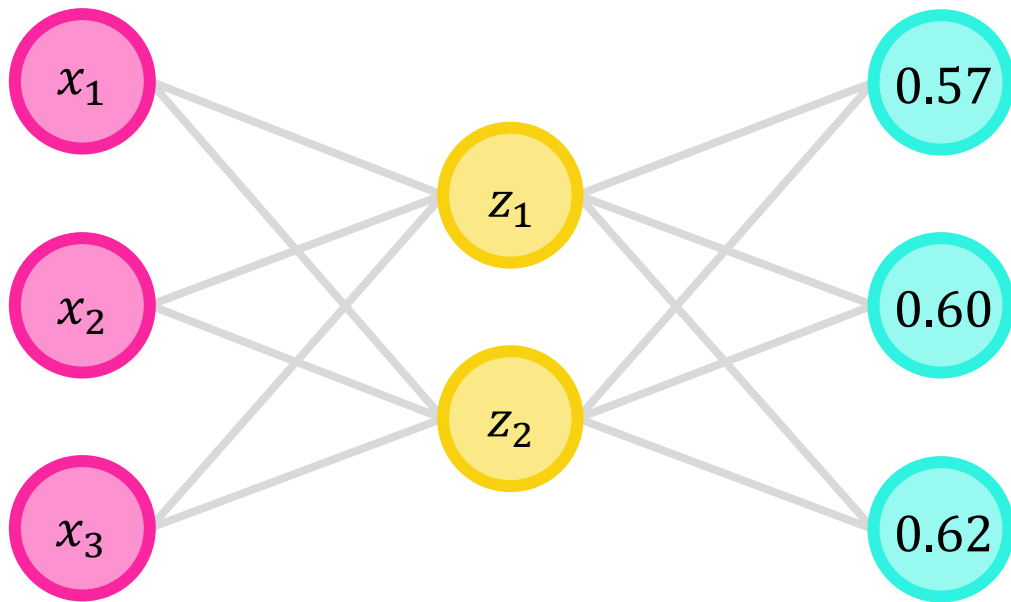
$$\text{인코더가중치: } W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

$$\text{디코더가중치: } W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$\partial L / \partial W_{encoder}$ 도 다음과 같이 전개할 수 있습니다.



$$\frac{\partial L}{\partial W_{encoder}} = \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}}$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

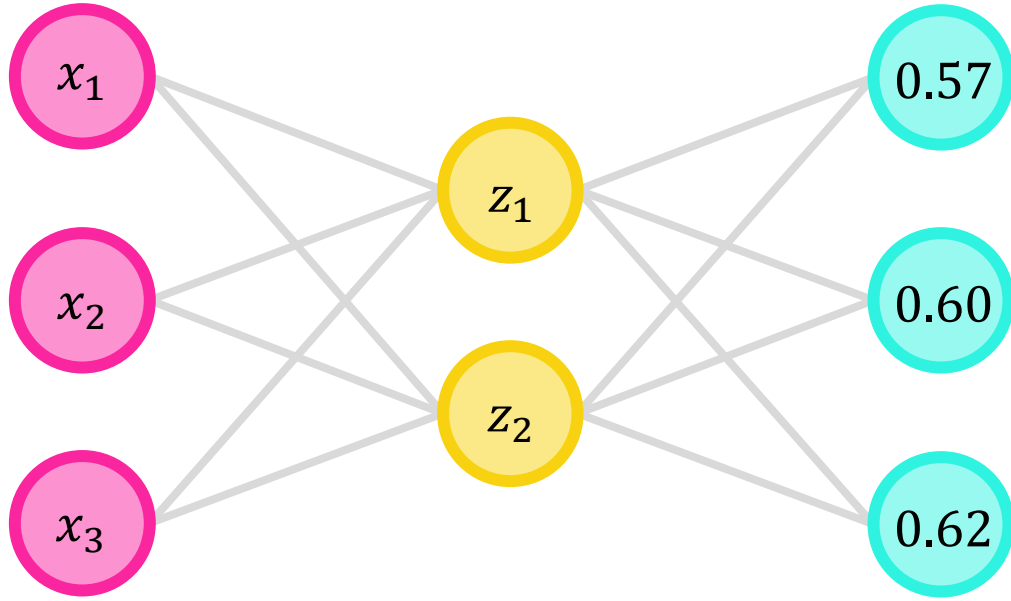
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

그리고 MSE의 공식에 의해서 $\partial L / \partial \hat{x}$ 는 다음과 같이 바뀌고,



$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

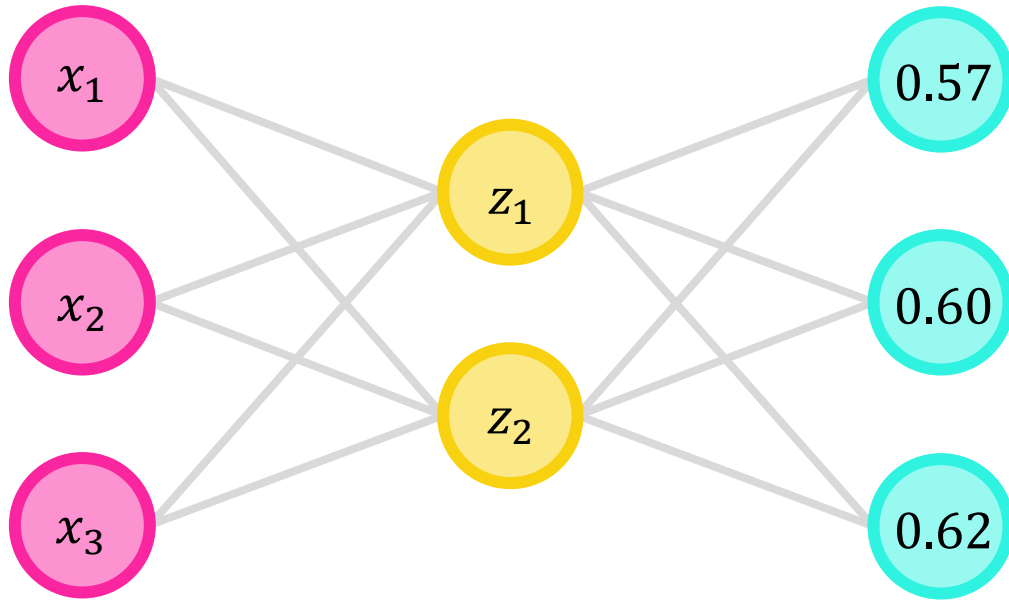
인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

그리고 $\partial \hat{x} / \partial z$ 는 다음 공식에 의해서



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

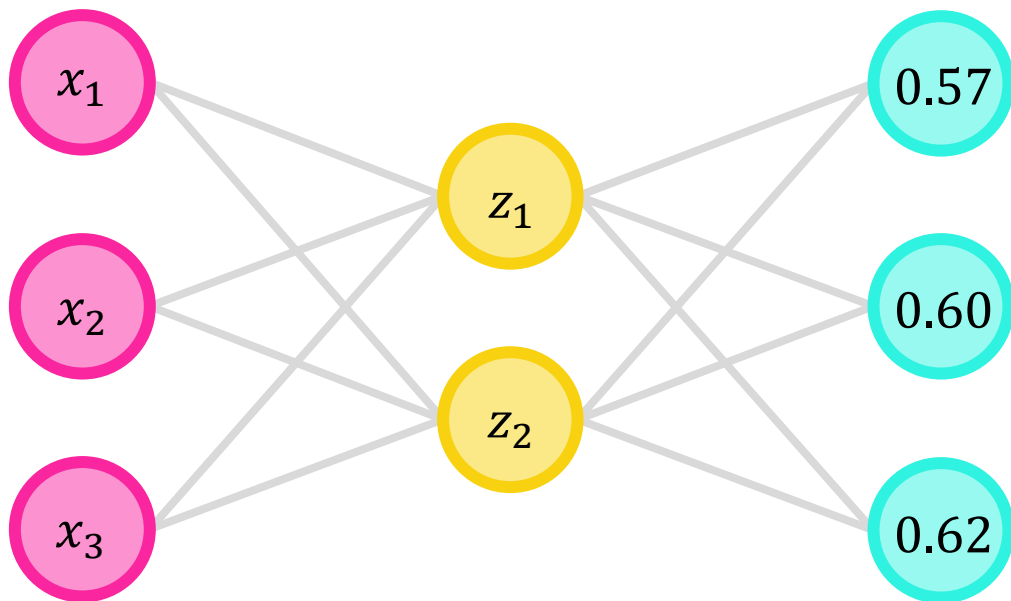
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \quad \hat{x} = \sigma(z \cdot W_{decoder}) \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

이렇게 전개할 수 있고,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

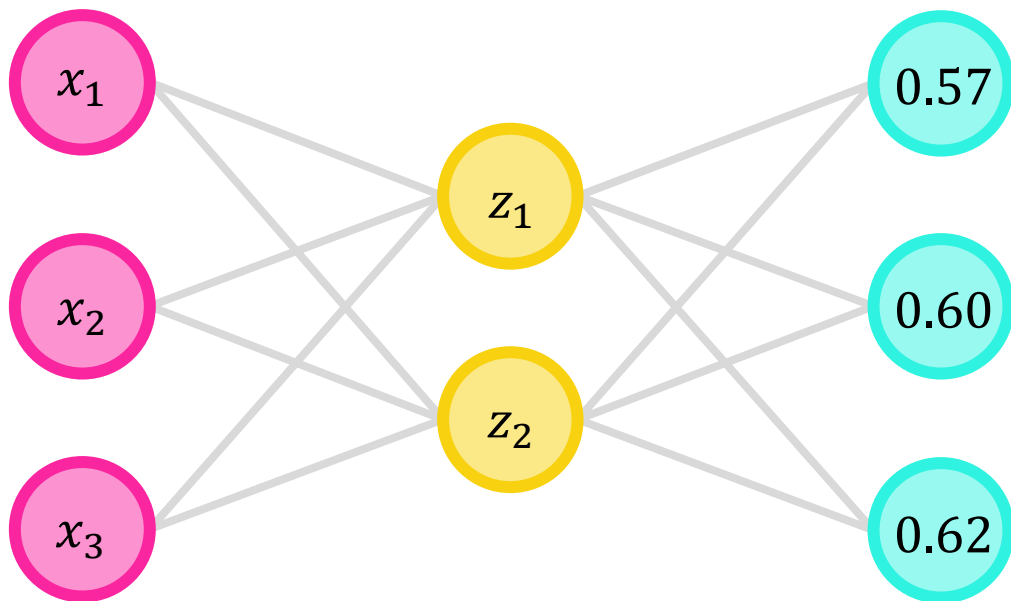
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$$\begin{aligned}
 \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \quad \hat{x} = \sigma(z \cdot W_{decoder}) \\
 &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\
 &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \sigma(z \cdot W_{decoder}) \cdot (1 - \sigma(z \cdot W_{decoder})) \cdot \frac{\partial z}{\partial W_{encoder}}
 \end{aligned}$$

다시 줄여서 표현하면, 이렇게 됩니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{\text{encoder}} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

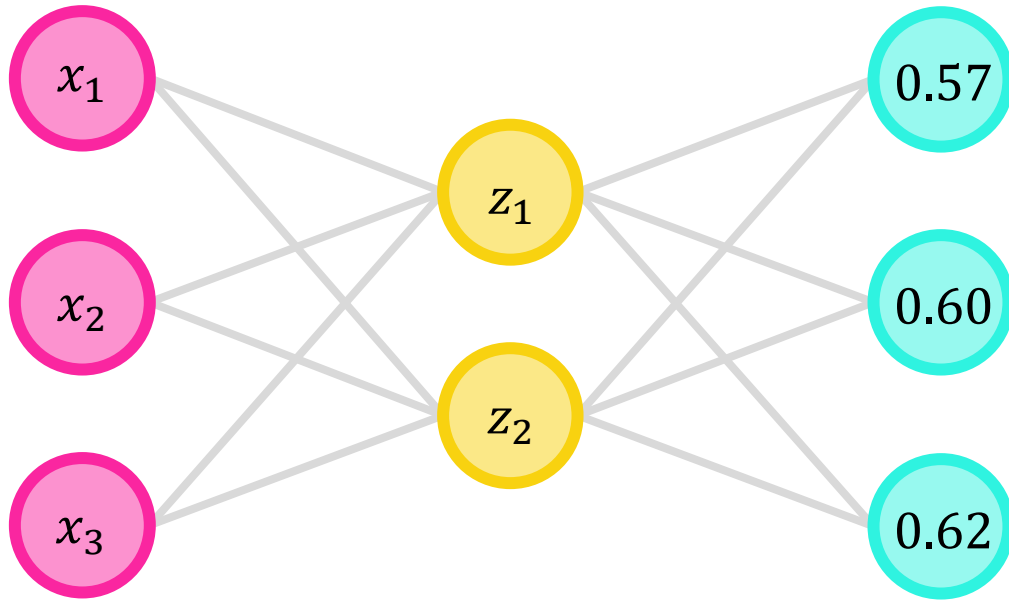
$$\text{디코더가중치: } W_{\text{decoder}} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{\text{decoder}}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{aligned} \frac{\partial L}{\partial W_{\text{encoder}}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{\text{encoder}}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{\text{encoder}}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{\text{decoder}}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot \frac{\partial z}{\partial W_{\text{encoder}}} \end{aligned}$$

그리고 이 부분은,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

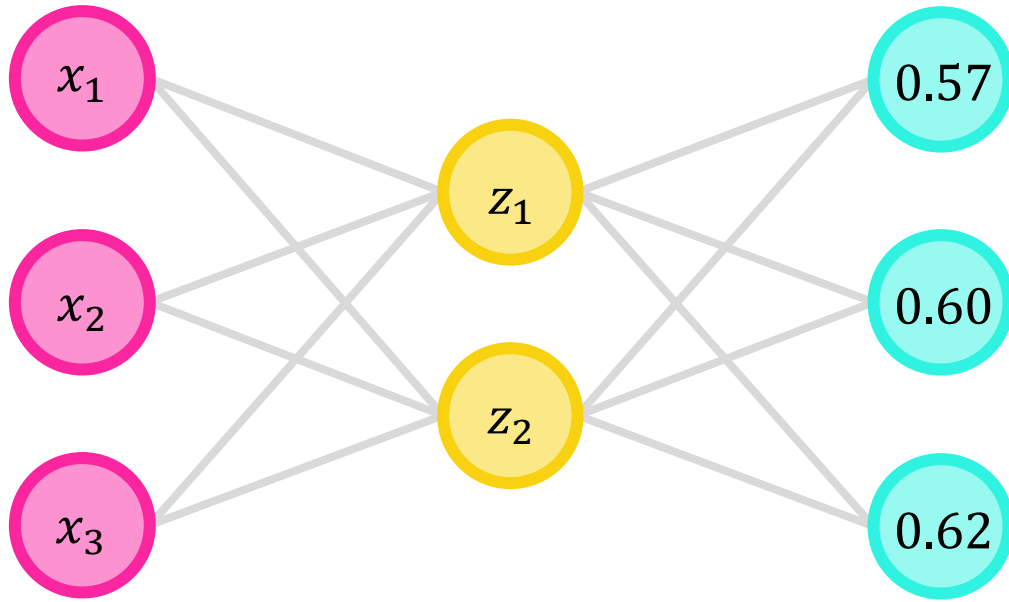
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

이 공식에 의해서,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

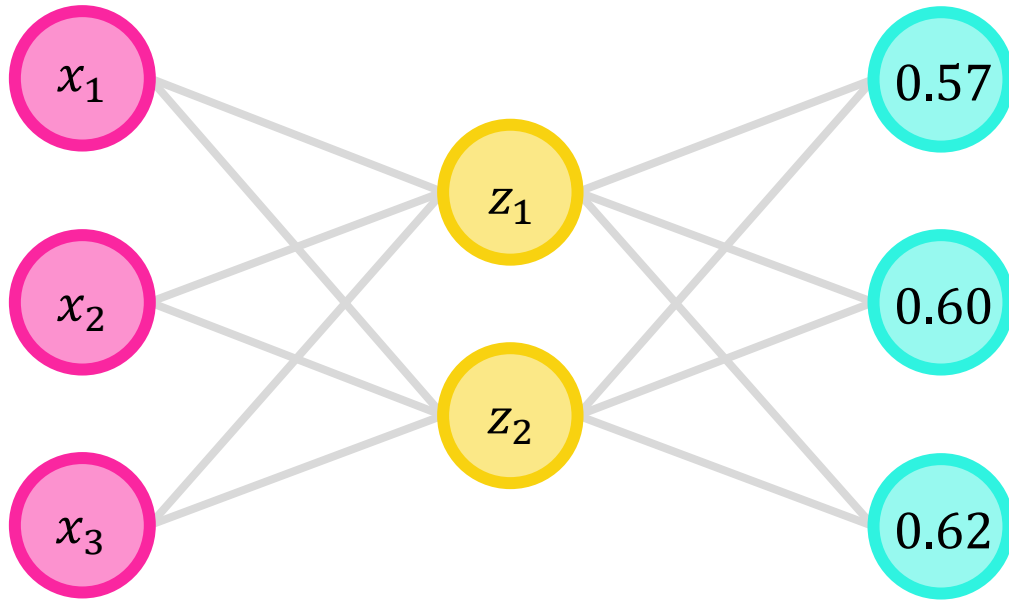
$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

$$= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot \frac{\partial z}{\partial W_{encoder}}$$

$z = \sigma(X \cdot W_{encoder})$

이렇게 전개할 수 있고,



$$\begin{aligned}\frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}}\end{aligned}$$

$$= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot x \cdot \sigma(X \cdot W_{encoder})(1 - \sigma(X \cdot W_{encoder}))$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

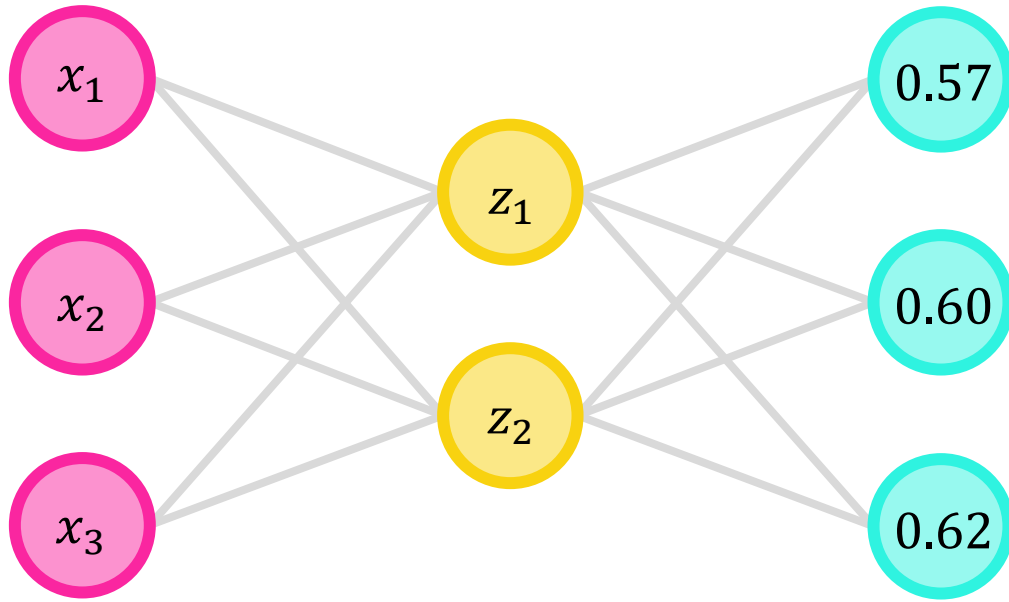
$$\text{디코더가중치: } W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\curvearrowright z = \sigma(X \cdot W_{encoder})$$

이렇게 줄여 쓸 수가 있습니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{\text{encoder}} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

$$\text{디코더가중치: } W_{\text{decoder}} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

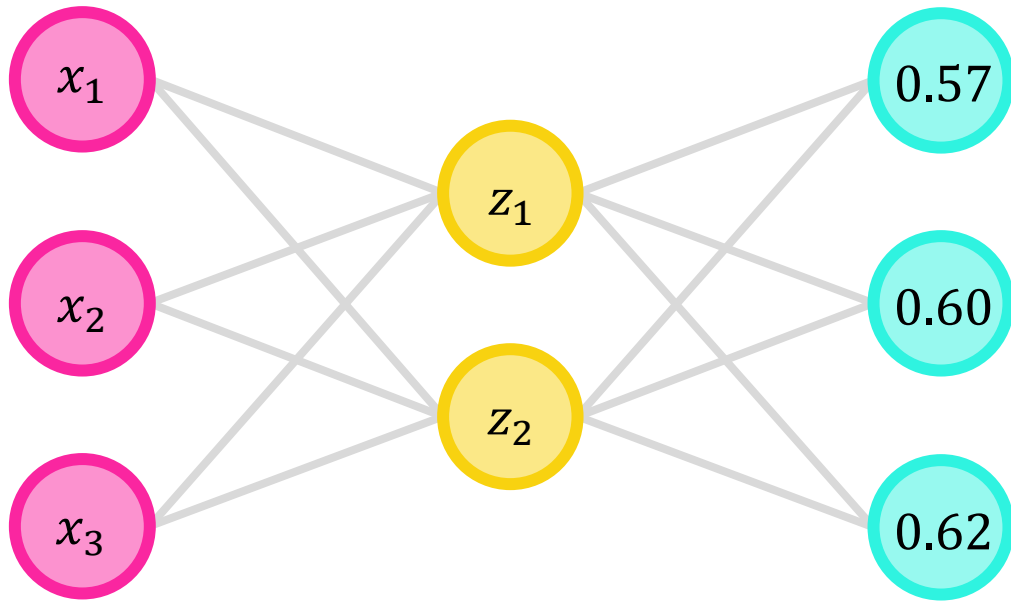
인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{\text{decoder}}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{aligned} \frac{\partial L}{\partial W_{\text{encoder}}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{\text{encoder}}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{\text{encoder}}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{\text{decoder}}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot x \cdot z \cdot (1 - z) \end{aligned}$$

$$z = \sigma(X \cdot W_{\text{encoder}})$$

그리고 값을 넣어서 계산하면, 다음과 같이 계산할 수 있습니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

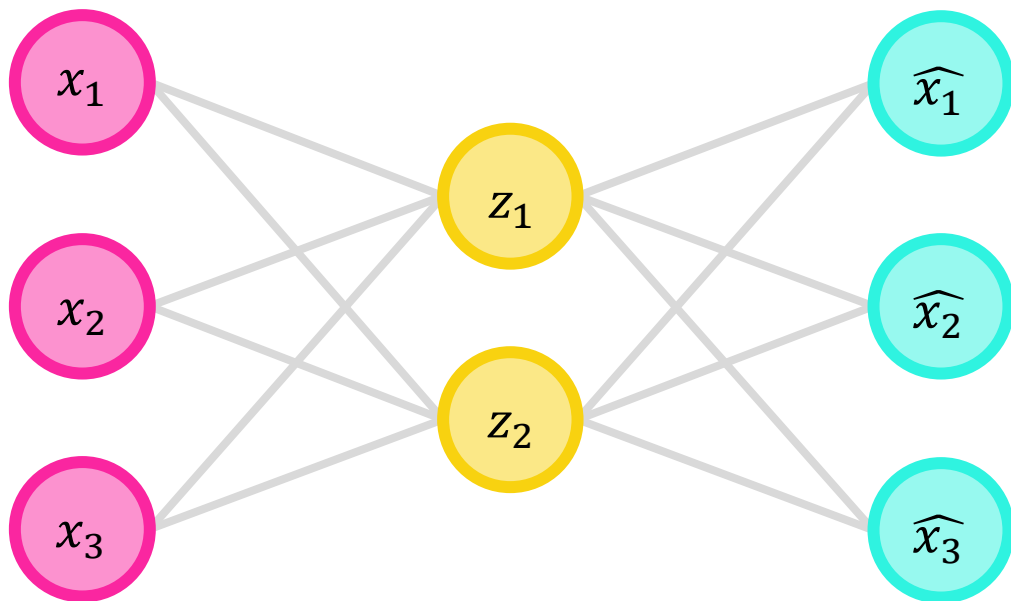
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot x \cdot z \cdot (1 - z) \\ &= \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix} \end{aligned}$$

이렇게 각각 가중치에 대한 gradient를 구하면,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

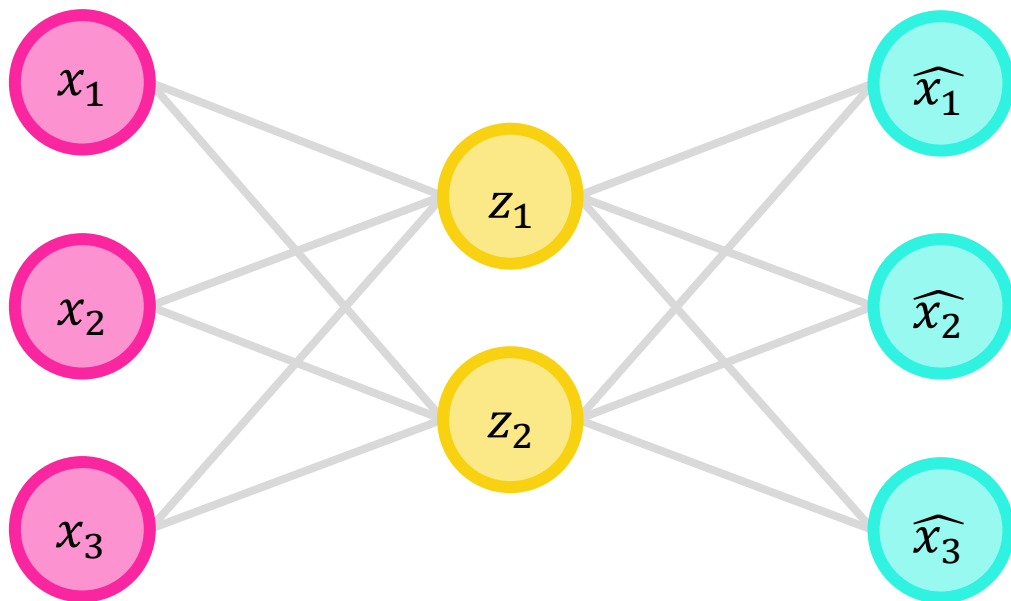
$$\text{디코더가중치: } W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

이렇게 경사하강법을 사용하여 가중치를 업데이트 할 수가 있습니다



$$W_{decoder}^* = W_{decoder} - \alpha \cdot \frac{\partial L}{\partial W_{decoder}}$$

$$W_{encoder}^* = W_{encoder} - \alpha \cdot \frac{\partial L}{\partial W_{encoder}}$$

입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

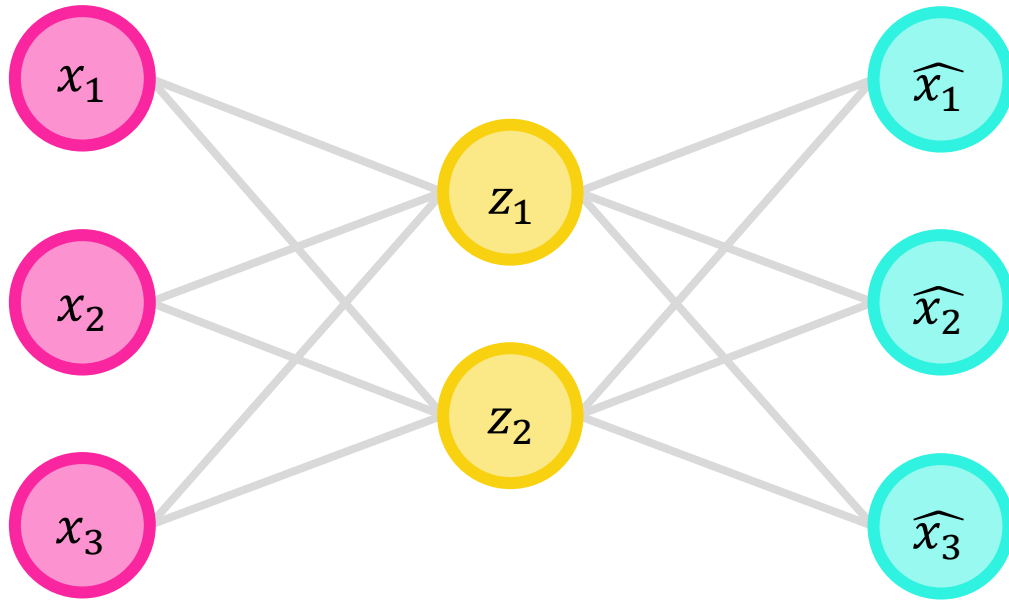
$$\text{디코더가중치: } W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

학습률 α 를 0.01로 하면,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

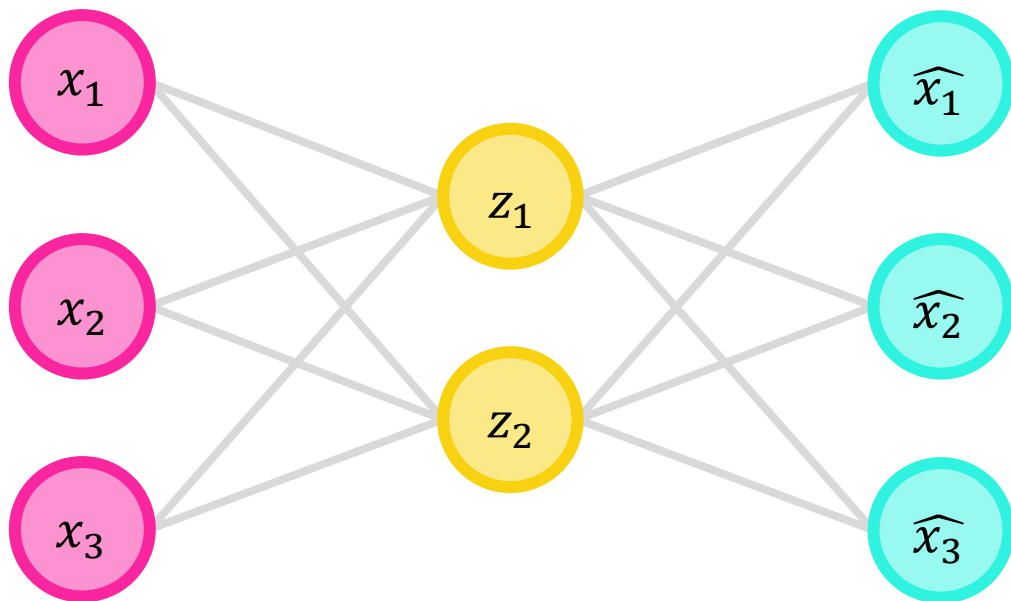
$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$

$$W_{decoder}^* = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$W_{encoder}^* = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

새로운 가중치는 다음과 같이 업데이트 됩니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

$$\text{인코더가중치: } W_{\text{encoder}} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$$

$$\text{디코더가중치: } W_{\text{decoder}} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

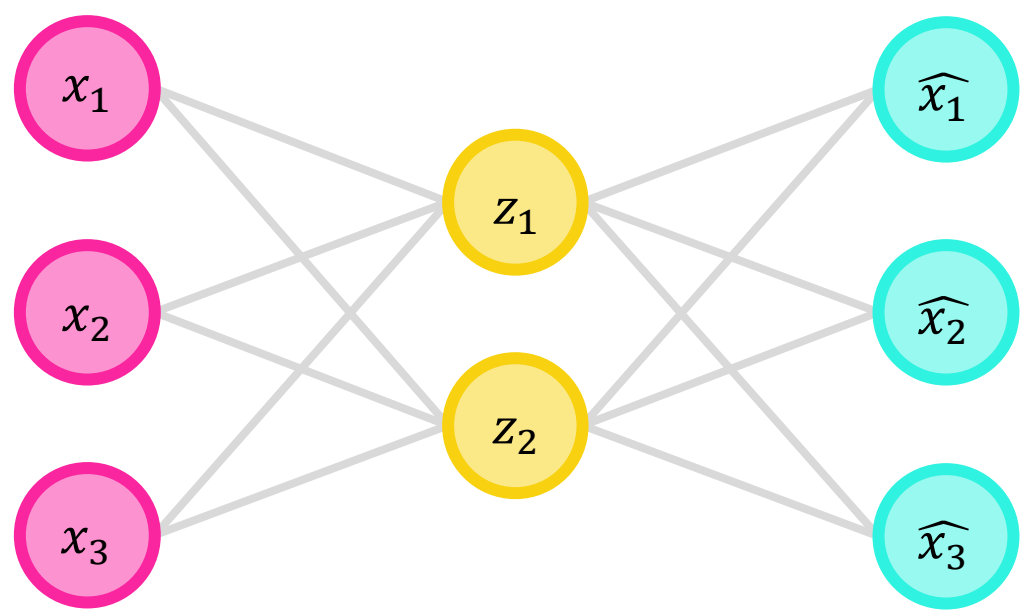
$$\frac{\partial L}{\partial W_{\text{decoder}}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\frac{\partial L}{\partial W_{\text{encoder}}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

$$\begin{bmatrix} 0.09872112 & 0.19893901 & 0.29915589 \\ 0.39868708 & 0.49891077 & 0.59913342 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{bmatrix} 0.09999431 & 0.19997739 \\ 0.29998111 & 0.39995311 \\ 0.49996618 & 0.59993286 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

이런 과정을 수없이 많이 반복하게 되면,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

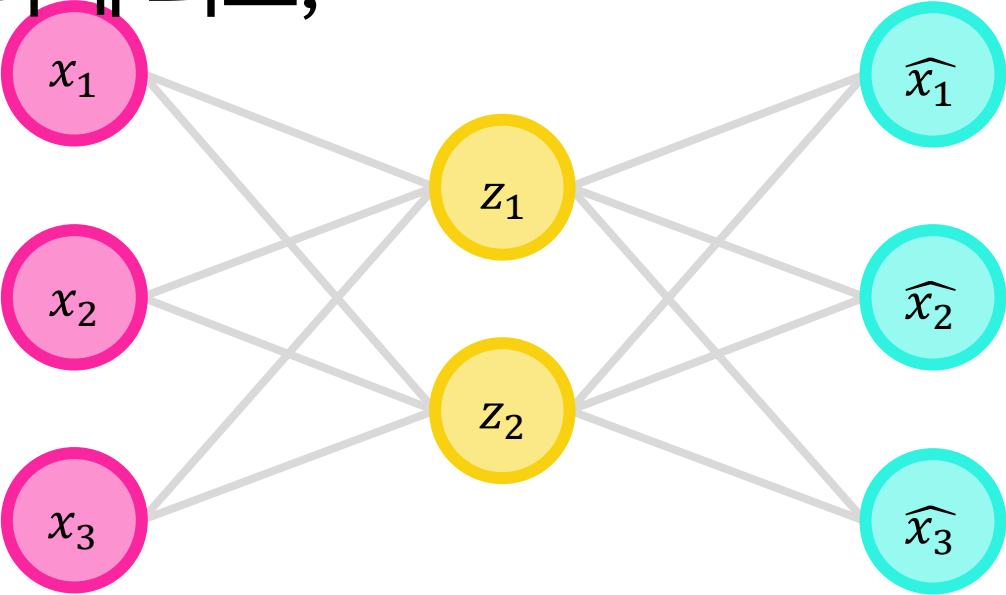
$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

$$\begin{bmatrix} 0.09872112 & 0.19893901 & 0.29915589 \\ 0.39868708 & 0.49891077 & 0.59913342 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{bmatrix} 0.09999431 & 0.19997739 \\ 0.29998111 & 0.39995311 \\ 0.49996618 & 0.59993286 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

잠재공간 z는 입력값의 핵심 특성을 표현하는 압축된 형태의 피쳐맵을 형성하게 되고,



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

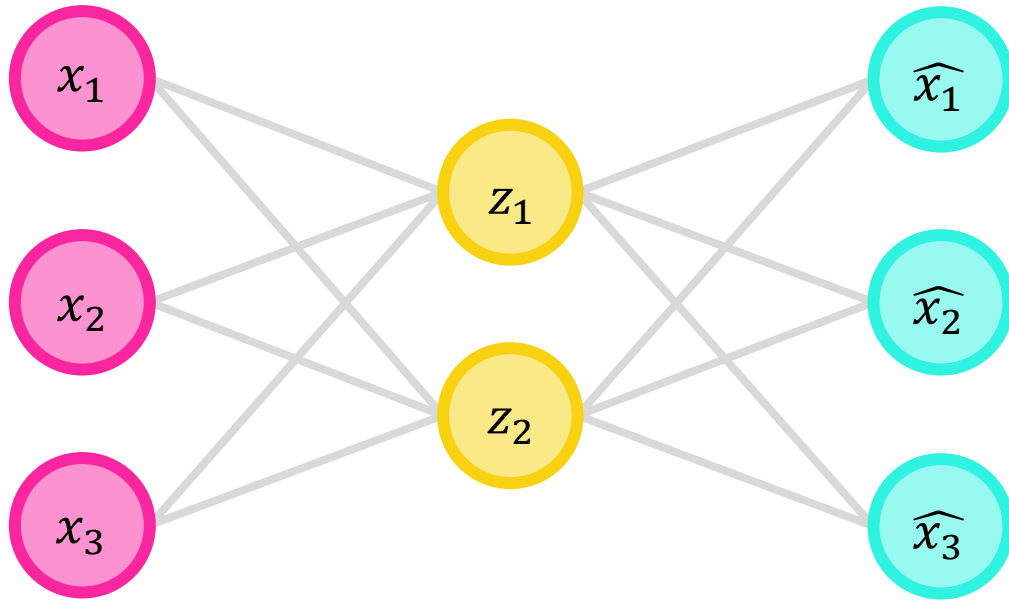
$$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

$$\begin{bmatrix} 0.09872112 & 0.19893901 & 0.29915589 \\ 0.39868708 & 0.49891077 & 0.59913342 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{bmatrix} 0.09999431 & 0.19997739 \\ 0.29998111 & 0.39995311 \\ 0.49996618 & 0.59993286 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

입력값과 거의 동일한 출력값을 내는 오토인코더 신경망이 될 것입니다.



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.55 \quad 0.57]$ 손실: 0.16

$\frac{\partial L}{\partial W_{decoder}} = \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$

$\frac{\partial L}{\partial W_{encoder}} = \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$

$$\begin{bmatrix} 0.09872112 & 0.19893901 & 0.29915589 \\ 0.39868708 & 0.49891077 & 0.59913342 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.13 & 0.11 & 0.08 \\ 0.13 & 0.11 & 0.09 \end{bmatrix}$$

$$\begin{bmatrix} 0.09999431 & 0.19997739 \\ 0.29998111 & 0.39995311 \\ 0.49996618 & 0.59993286 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.0006 & 0.0023 \\ 0.0019 & 0.0047 \\ 0.0034 & 0.0067 \end{bmatrix}$$

마지막으로, 오토인코더의 활용에 대해 말씀드리고자 합니다

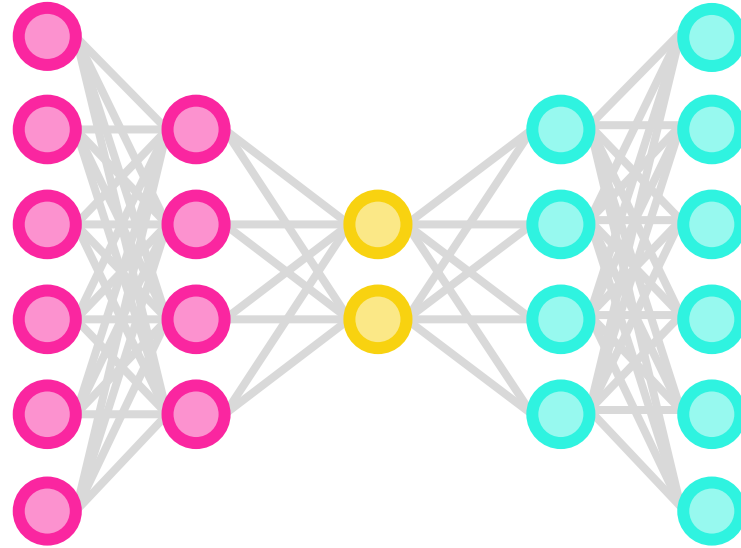
마지막으로, 오토인코더의 활용에 대해 말씀드리고자 합니다

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.



오토인코더는 다음과 같은 주요한 특성으로 인해

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.

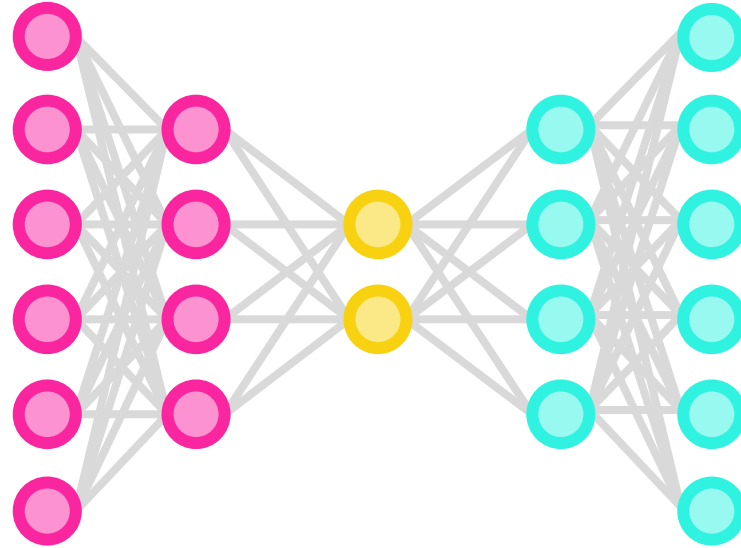


예를들면, 스캔된 문서 이미지의 노이즈를 제거한다든지,

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.



There exist several methods to design forms with fields to fields may be surrounded by bounding boxes, by light rectangles o methods specify where to write and, therefore, minimize the effe with other parts of the form. These guides can be located on a s is located below the form or they can be printed directly on the fo a separate sheet is much better from the point of view of the que but requires giving more instructions and, more importantly, rest this type of acquisition is used. Guiding rulers printed on the used for this reason. Light rectangles can be removed more easily whenever the handwritten text touches the rulers. Nevertheless, be taken into account: The best way to print these light rectan.



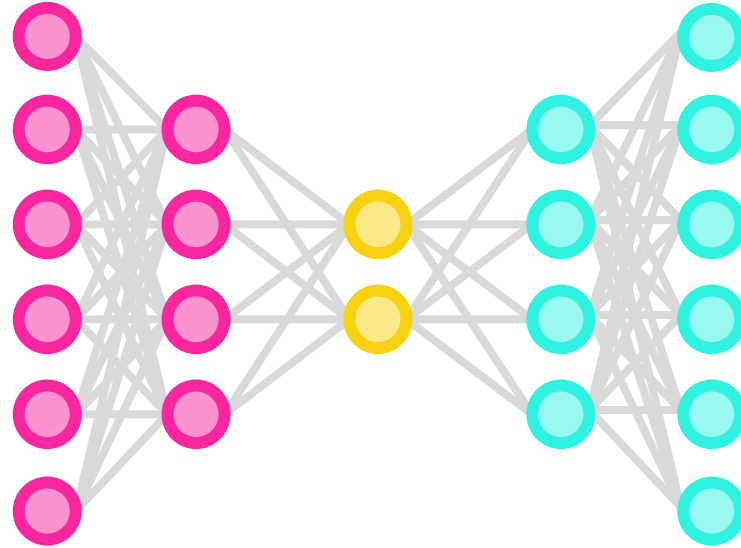
There exist several methods to design forms with fields to fields may be surrounded by bounding boxes, by light rectangles o methods specify where to write and, therefore, minimize the effe with other parts of the form. These guides can be located on a s is located below the form or they can be printed directly on the fo a separate sheet is much better from the point of view of the que but requires giving more instructions and, more importantly, rest this type of acquisition is used. Guiding rulers printed on the used for this reason. Light rectangles can be removed more easily whenever the handwritten text touches the rulers. Nevertheless, be taken into account: The best way to print these light rectan.

예를들면, 스캔된 문서 이미지의 노이즈를 제거한다든지,

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.



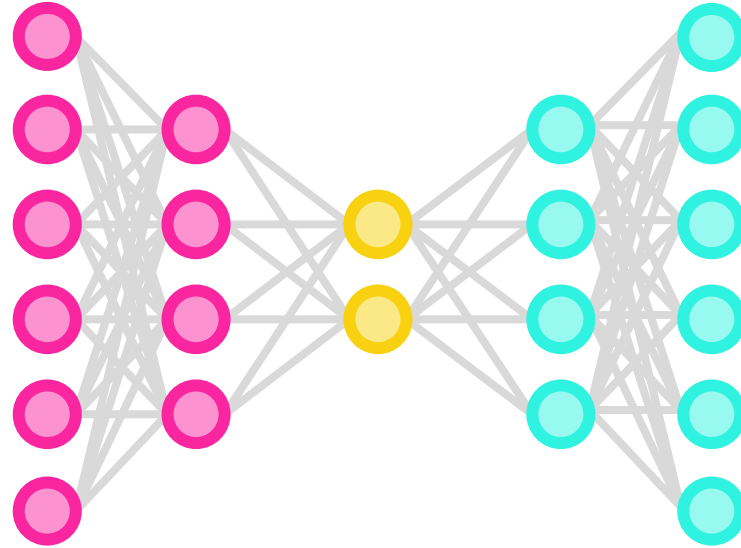
There exist several methods to design forms with fields to fields may be surrounded by bounding boxes, by light rectangles o methods specify where to write and, therefore, minimize the effe with other parts of the form. These guides can be located on a s is located below the form or they can be printed directly on the fo a separate sheet is much better from the point of view of the que but requires giving more instructions and, more importantly, rest this type of acquisition is used. Guiding rulers printed on the used for this reason. Light rectangles can be removed more easily whenever the handwritten text touches the rulers. Nevertheless, be taken into account: The best way to print these light rectan.



There exist several methods to design forms with fields to fields may be surrounded by bounding boxes, by light rectangles o methods specify where to write and, therefore, minimize the effe with other parts of the form. These guides can be located on a s is located below the form or they can be printed directly on the fo a separate sheet is much better from the point of view of the que but requires giving more instructions and, more importantly, rest this type of acquisition is used. Guiding rulers printed on the used for this reason. Light rectangles can be removed more easily whenever the handwritten text touches the rulers. Nevertheless, be taken into account: The best way to print these light rectan.

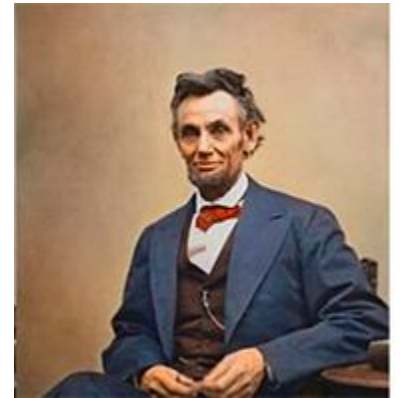
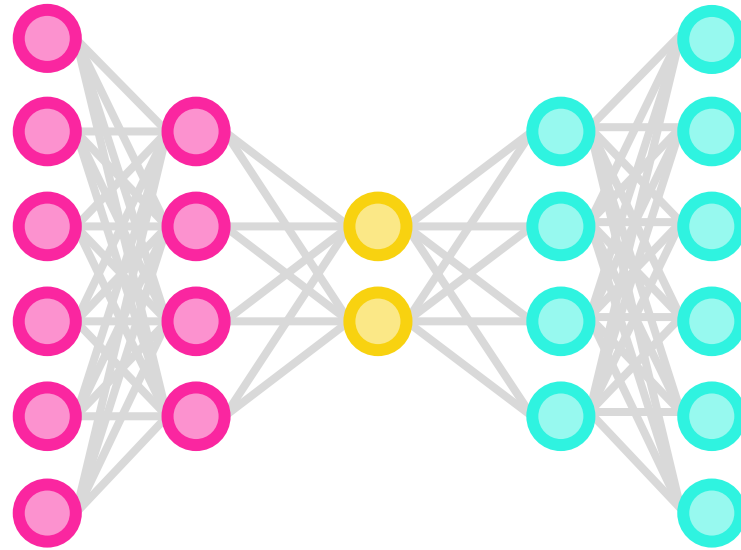
디지털 이미지나 영상의 노이즈를 제거한다든지,

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.



흑백 이미지의 컬러를 복원하는데 사용되는 등 여러가지 쓰임새가 많은 신경망입니다

오토인코더는 데이터의 차원축소, 그리고 축소된 잠재공간으로 부터의 데이터 생성기능으로 인해 1)중요 특성 학습, 2)이미지 노이즈 제거, 그리고 3)데이터 복원 등과 같은 일에 사용됩니다.



자 여기까지가 오늘 제가 준비한
오토인코더에 관한 소개 영상입니다

다음시간에도 쉽고 재미있는 딥러닝
영상으로 찾아뵙겠습니다

간단하면서도 활용도가 높은 오토인코더를
많이 사랑해주세요!



감사합니다!

좋은 하루 되세요!!

이 채널은 여러분의 관심과 사랑이 필요합니다

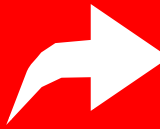
좋아요



댓글



공유



구독



‘좋아요’와 ‘구독’버튼은 강의 준비에 큰 힘이 됩니다!

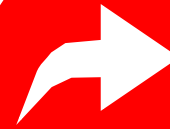
좋아요



댓글



공유



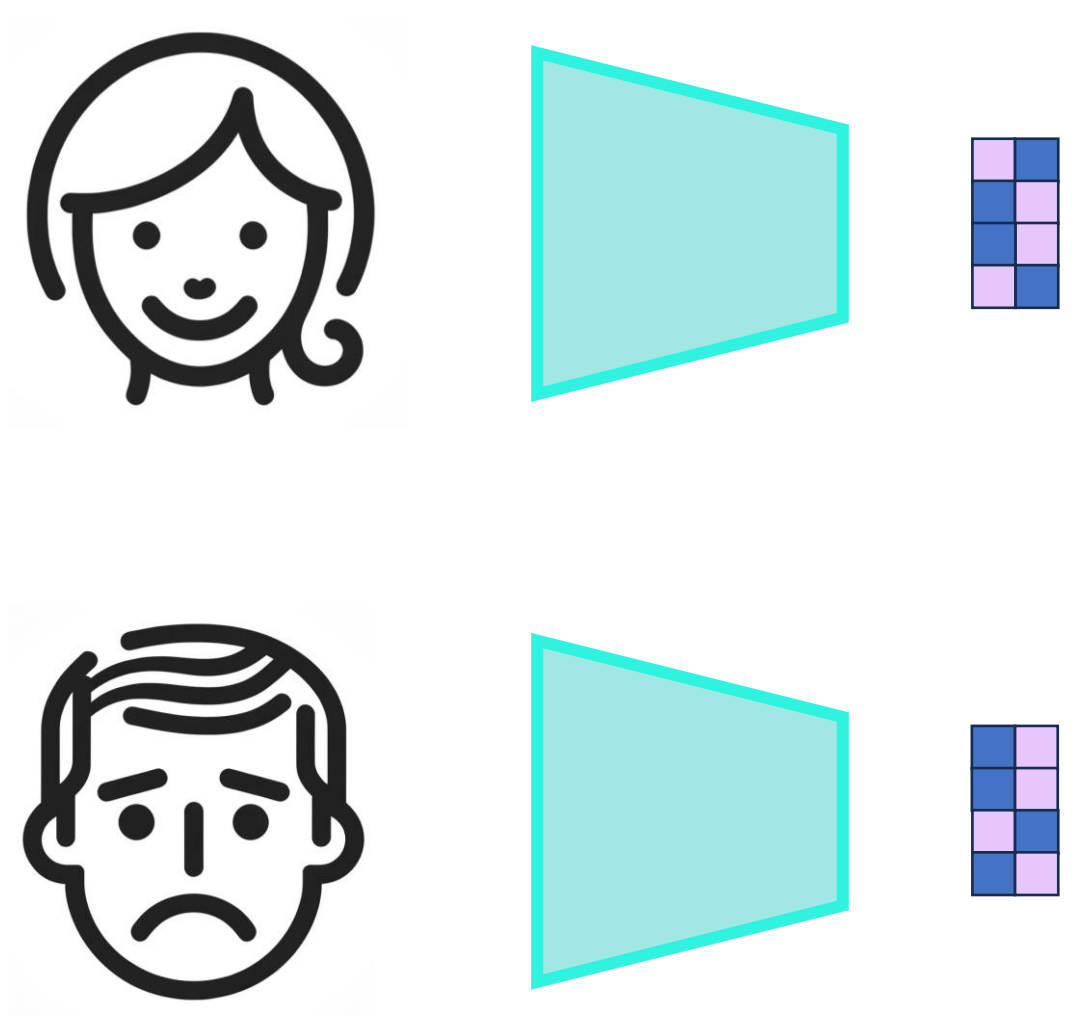
구독



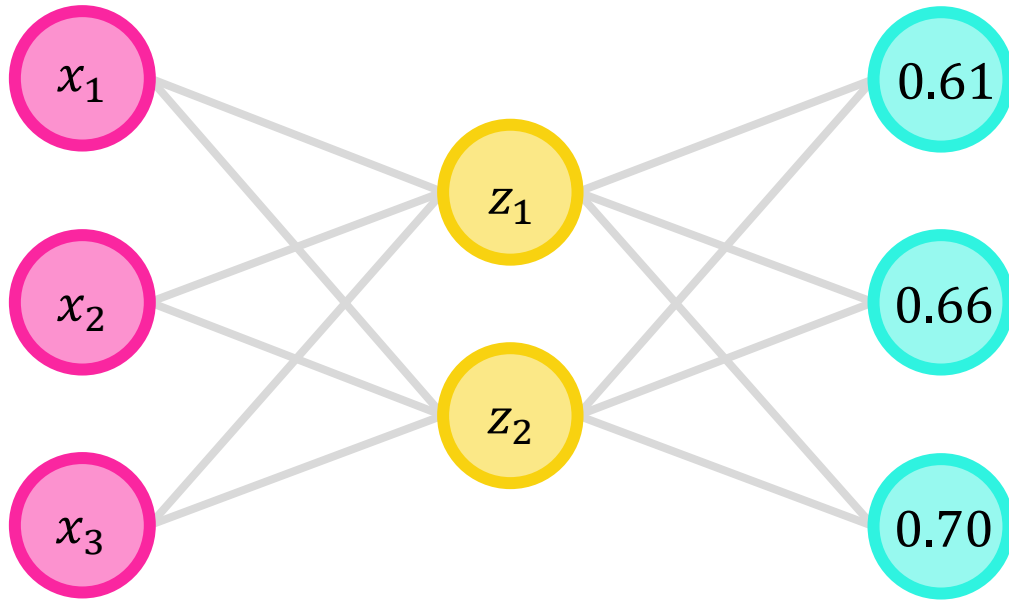
그리고 영상 자료를 사용하실때는
출처 '신박AI'를 밝혀주세요



그래서 다음과 같은 활용법을 예로 들어 설명하고자 합니다



$\partial z / \partial W_{encoder}$ 은 다음의 공식에 의해서



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

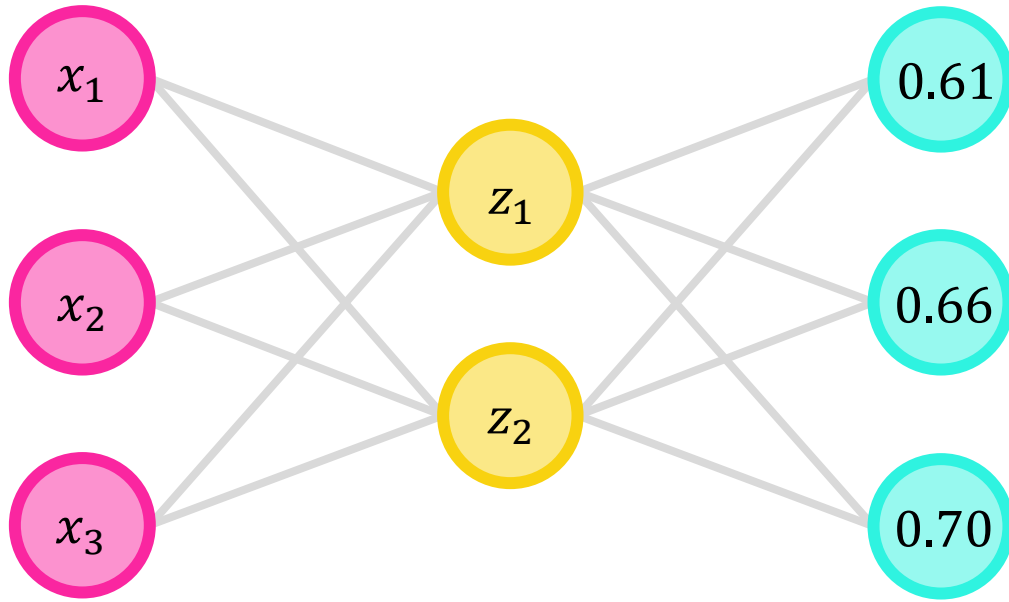
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.9 \quad 0.94]$ 손실: 2.42

$$z = \sigma(x \cdot W_{encoder})$$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

$\partial z / \partial W_{encoder}$ 은 다음의 공식에 의해서



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

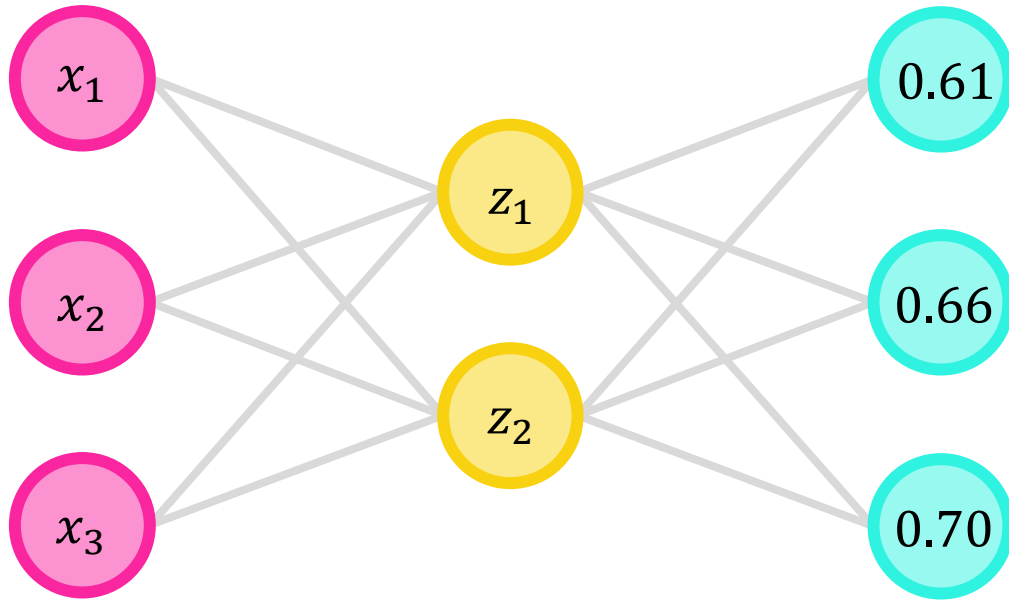
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.9 \quad 0.94]$ 손실: 2.42

$$z = \sigma(x \cdot W_{encoder})$$

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot \frac{\partial z}{\partial W_{encoder}} \end{aligned}$$

이렇게 바뀌게 됩니다



입력값: $x = (0.1, 0.2, 0.3)$ σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

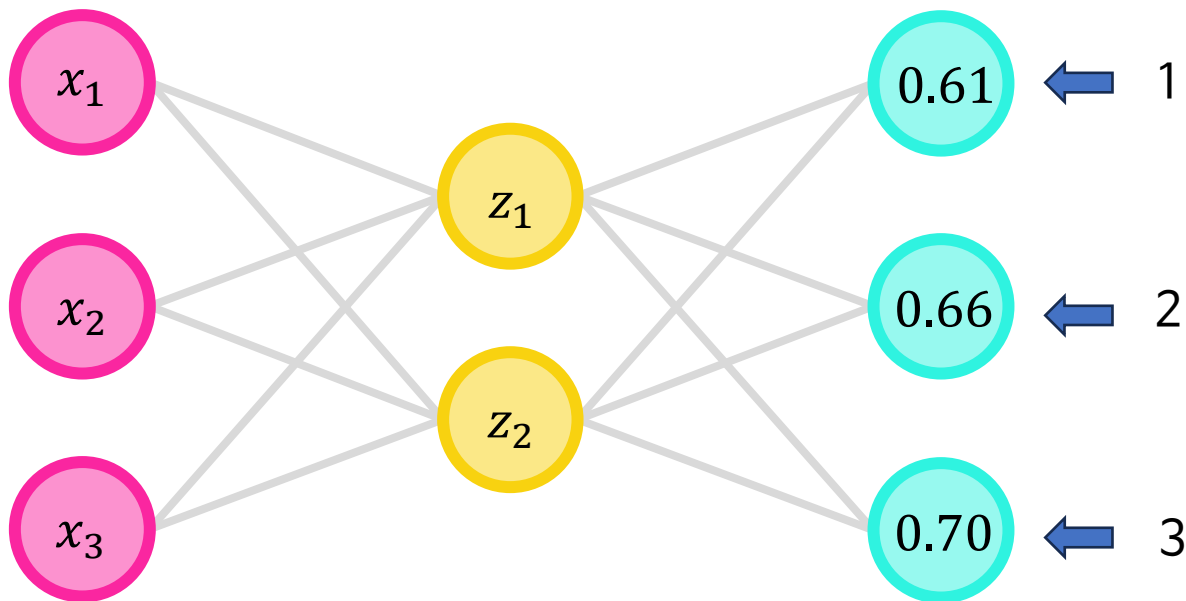
디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $z = [0.9 \quad 0.94]$ 손실: 2.42

$$\begin{aligned} \frac{\partial L}{\partial W_{encoder}} &= \frac{\partial L}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} = 2 \cdot (\hat{x} - x) \cdot \frac{\partial \hat{x}}{\partial z} \cdot \frac{\partial z}{\partial W_{encoder}} \\ &= 2 \cdot (\hat{x} - x) \cdot W_{decoder}^T \cdot \hat{x} \cdot (1 - \hat{x}) \cdot x^T \cdot z \cdot (1 - z) \end{aligned}$$

$z = \sigma(x \cdot W_{encoder})$

손실 loss는 2.42가 나왔습니다



입력값: $x = (1, 2, 3)$

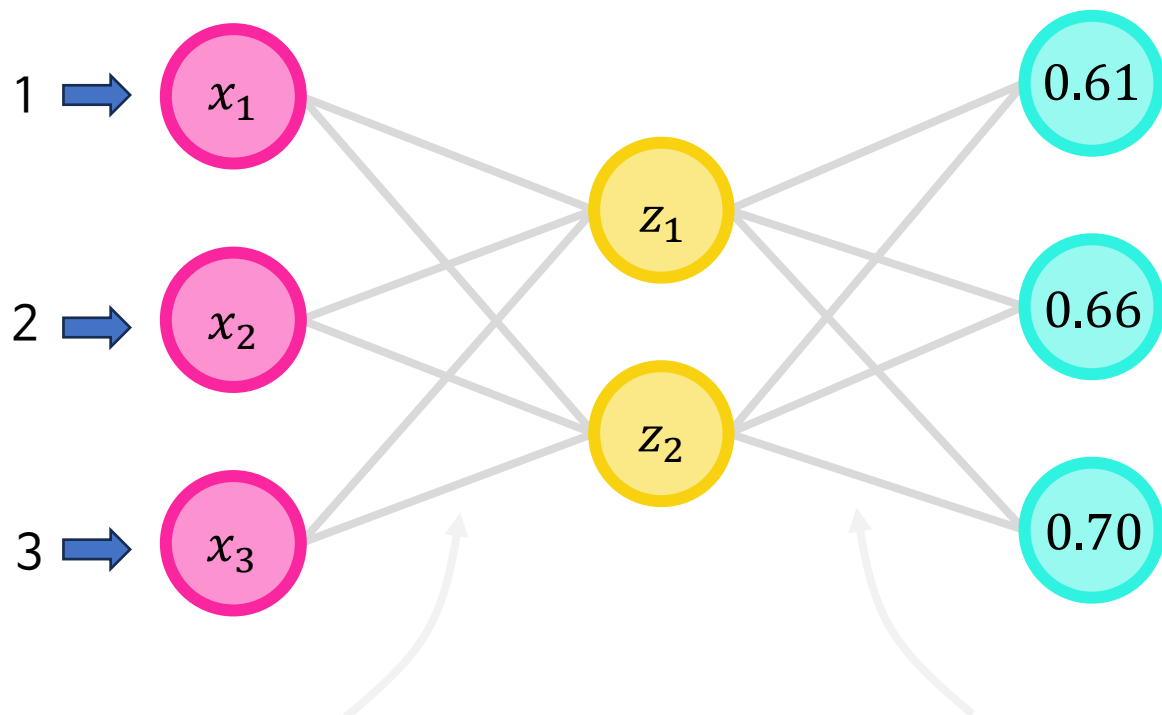
σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

인코딩: $Z = [0.9 \quad 0.94]$ 손실: 2.42

오토인코더의 손실 계산법은 너무너무 간단합니다



입력값: $x = (1, 2, 3)$

σ : 시그모이드

인코더가중치: $W_{encoder} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$

디코더가중치: $W_{decoder} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}$

$$\begin{aligned} Z &= \sigma(X \cdot W_{encoder}) \\ Z &= \sigma([1 \ 2 \ 3] \cdot \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}) \\ Z &= \sigma([2.2 \ 2.8]) \\ Z &= [0.9 \ 0.94] \end{aligned} \quad \begin{aligned} \hat{X} &= \sigma(Z \cdot W_{decoder}) = \sigma([0.9 \ 0.94] \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix}) \\ &= \sigma([0.47 \ 0.65 \ 0.84]) \\ &= [0.61 \ 0.66 \ 0.70] \end{aligned}$$

지금까지 시청해 주셔서 감사합니다

다음시간에는 더 유익하고 재미있는 영상으
로 찾아뵙겠습니다

그때까지 안녕히 계세요

감사합니다!

좋은 하루 되세요!!

이 채널은 여러분의 관심과 사랑이 필요합니다

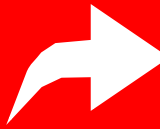
좋아요



댓글



공유



구독



‘좋아요’와 ‘구독’버튼은 강의 준비에 큰 힘이 됩니다!

좋아요



댓글



공유



구독



그리고 영상 자료를 사용하실때는
출처 '신박AI'를 밝혀주세요





Copyright © 2024 by 신박AI

All rights reserved

본 문서(PDF)에 포함된 모든 내용과 자료는 저작권법에 의해 보호받고 있으며, 신박AI에 의해 제작되었습니다.

본 자료는 오직 개인적 학습 목적과 교육 기관 내에서의 교육용으로만 무료로 제공됩니다.

이를 위해, 사용자는 자료 내용의 출처를 명확히 밝히고,

원본 내용을 변경하지 않는 조건 하에 본 자료를 사용할 수 있습니다.

상업적 사용, 수정, 재배포, 또는 이 자료를 기반으로 한 2차적 저작물 생성은 엄격히 금지됩니다.

또한, 본 자료를 다른 유튜브 채널이나 어떠한 온라인 플랫폼에서도 무단으로 사용하는 것은 허용되지 않습니다.

본 자료의 어떠한 부분도 상업적 목적으로 사용하거나 다른 매체에 재배포하기 위해서는 신박AI의 명시적인 서면 동의가 필요합니다.

위의 조건들을 위반할 경우, 저작권법에 따른 법적 조치가 취해질 수 있음을 알려드립니다.

본 고지 사항에 동의하지 않는 경우, 본 문서의 사용을 즉시 중단해 주시기 바랍니다.

