* Types of autoencoder :- (1)

(1) UnderComplete Autoencoder :-

- The objective of undercomplete autoencoder is to capture the most important features present in the data.
- That means it reads only important features in given data.
- undercomplete autoencoders have smaller dimensions for hidden layer compared to input layer. This helps to obtain important features from data.
  means it have smaller insize hidden layer than input layer because it helps to btain or focus on important data. (Hidden layer has fewer neurons than the input layer.
- It minimizes the loss function by penalizing the $g(f(x))$ for being different from input x.
  It uses MSE for reconstruction.

* Structure of Autoencoder :-

The network Consist of :-

1. input layer(x) :- The original input data.
2. Encoder layer :- reduces dimentidity & extract important features.
3. Bottleneck layer(z) :- The compressed, low dimentional representation.
4. Decoder layer :- Expands the compressed data back to it's original forms.
5. Output layer($\hat{x}$) :- The reconstructed version of input.

Autoencoder :-

Mathematical Representation :-

1. Encoding function :-

$$Z = f(x) = \sigma(Wx + b)$$

Where,

X = input vector. W = weight matrix (Encoder)

b = bias Z = Encoded representation.

σ = Activation function (Relu, Sigmoid, Tanh)

2. Decoding function :-

$$\hat{X} = g(z) = \sigma(W'z + b')$$

Where,

W' = Decoder weight matrix

b' = Decoder bias

$\hat{X}$ = Reconstructed input.

Loss function MSE :-

$$MSE = \frac{1}{N}\Sigma(x - \hat{x})^2$$

example :- Define input data :-

Consider a simple 2D input vector :

$$X = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Step 1. Define encoder parameter

Encoder weight matrix W & bias b :

$$W = [0.5 \quad 0.3] \qquad b = [0.2]$$

compute encoded representation $z := Wx + b$

(3)

$$[0.5 \quad 0.3] \begin{bmatrix} 2 \\ 3 \end{bmatrix} = (0.5 \times 2) + (0.3 \times 3)$$

$$= 1.0 + 0.9 = 1.9$$

$Z = 1.9 + 0.2$ ← added bias

$X = 2.1$

$Z = 1.9 + 0.2$

$= 2.1$

So, Compressed representation :- $Z = [2.1]$

Step 2 : Define Decoder Parameter :-

The decoder reconstruct x from z.

$$W' = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \& \; b' = \begin{bmatrix} 0.1 \\ 0.05 \end{bmatrix}$$

reconstructed output :-

$$\hat{x} = W'z + b'$$

$$\hat{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \times (2.1) = \begin{bmatrix} (0.4 \times 2.1) \\ (0.6 \times 2.1) \end{bmatrix} = \begin{bmatrix} 0.84 \\ 1.26 \end{bmatrix}$$

Now, add bias b'

$$\hat{x} = \begin{bmatrix} 0.84 + 0.1 \\ 1.26 + 0.05 \end{bmatrix} = \begin{bmatrix} 0.94 \\ 1.31 \end{bmatrix}$$

reconstructed output is $\hat{x} = \begin{bmatrix} 0.94 \\ 1.31 \end{bmatrix}$

Step 3 :- Compute MSE.

$$MSE = \frac{1}{N} \Sigma [x - \hat{x}]^2$$

$$= \frac{1}{2} \Sigma \left[(2 - 0.94)^2 + (3 - 1.31)^2 \right]$$

$$= \frac{1}{2} \left[(1.06)^2 + (1.69)^2 \right]$$

:-

$\frac{1}{2}[1.1236 + 2.8561]$ (4)

$= \frac{1}{2} \times 3.9797$

MSE $= 1.9899$.

---

* Denoising Autoencoder :-

It learn to remove noise from data.

A standard auto encoder learns to compress & reconstruct input without noise. I + add.

Here :-

* compression (Encoding)

The input $x = [2, 3]$ is originally 2-Dimensional

The encoder reduces it to 1D i.e. $z = 2.1$

This means you are forcing autoencoder to store the important information from 2 numbers to 1 number

* Reconstruction (Decoding)

. The decoder tries to rebuild original 2D data from $z = 2.1$

. It gives us :- $\hat{x} = [0.94, 1.31]$ which is close to $[2, 3]$ but not exactly same.