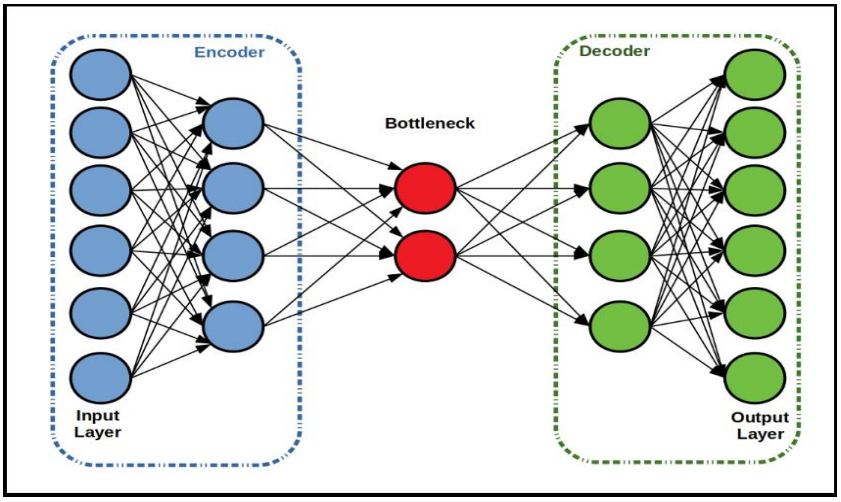
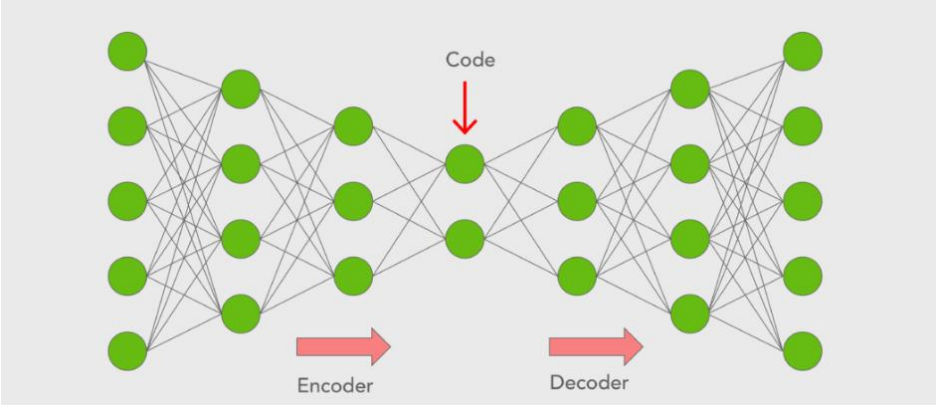
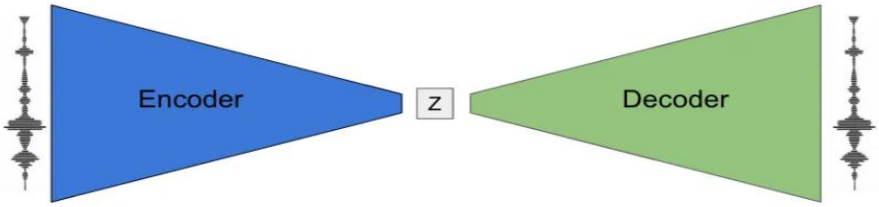
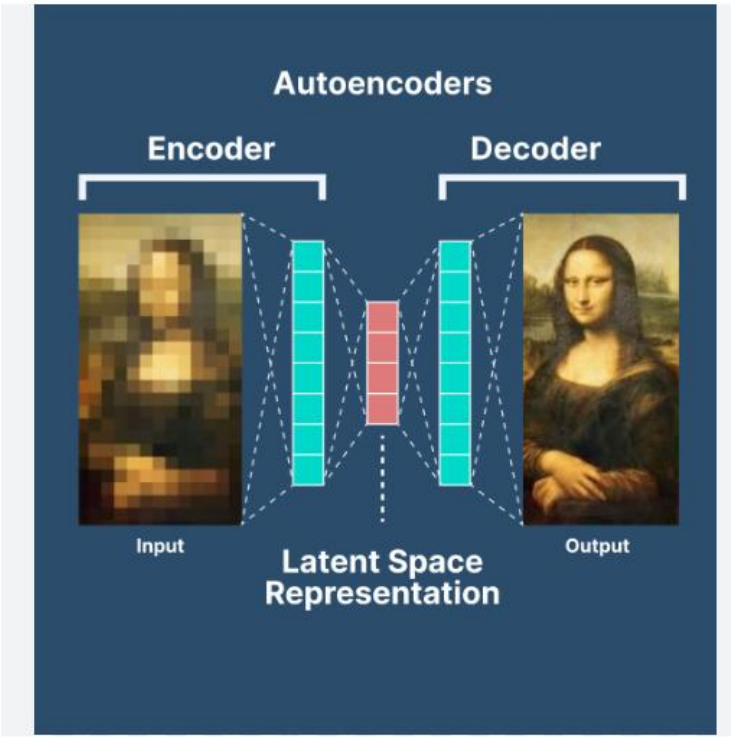
Autoencoder

- Autoencoder is a type of neural network where the output layer has the same dimensionality as the input layer.
- In simpler words, the number of output units in the output layer is equal to the number of input units in the input layer.
- An autoencoder replicates the data from the input to the output in an unsupervised manner and is therefore sometimes referred to as a replicator neural network.
- They work by compressing the input into a **latent-space representation** also known as **bottleneck/Code**, and then reconstructing the output from this representation.
- Autoencoder is an unsupervised algorithm.
- We can define autoencoder as **feature extraction algorithm**.









An Autoencoder consists of three layers:

1. Encoder
2. Code / Bottle neck
3. Decoder

- The Encoder layer compresses the input image into a latent space representation. It encodes the input image as a compressed representation in a reduced dimension.
- The compressed image is a distorted version of the original image.

- The Code layer represents the compressed input fed to the decoder layer.
- The decoder layer decodes the encoded image back to the original dimension. The decoded image is reconstructed from latent space representation, and it is reconstructed from the latent space representation and is a lossy reconstruction of the original image.

Application

Image Denoising

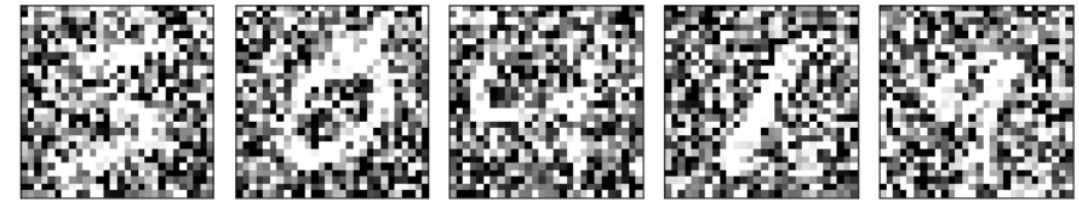
Dimensionality Reduction

Feature Extraction

Image Generation

Image colorization

Before adding noise



after adding noise