

What are Autoencoders?

Autoencoders are a type of neural network that can be thought of as a "compression and decompression" system. They are often used for tasks like image or data compression, but they can also be helpful in understanding neural networks.

Process of Autoencoders:

Compression: By taking a lot of data, such as pictures of cats, autoencoders take the data and then compress it into smaller representations. Almost like taking a big image and then squishing it into a tiny version while keeping the important details.

Encoding: During this compression phase, the autoencoder has an "encoder" which takes the input data and maps it into smaller sets of numbers. These numbers represent the important parts of the data in a compact form. This is described as the "latent representation."

Decompression: During this phase, the autoencoder will take the small set of numbers and decompress it to try and recreate the original data, (like a cat picture), almost like refactoring it to its original size.

Decoding: This decompression stage is called the decoder, where it takes the compressed number to generate an output where it's like the original data.

Training: Autoencoders are trained with the same process of backpropagation and gradient descent. The goal is to make sure that the decompressed output closely resembles the input. Basically, the network learns to compress and decompress data effectively.

Importance of Autoencoders:

The importance of autoencoders is obviously its ability to take large amounts of data and compress it into smaller, usable data, keeping in mind to maintain its important features. This effectively reduces data size (pretty much what we're aiming to do)

Autoencoders are also good with anomaly detection as they're already recognizing a normal in the data patterns, if something doesn't exactly compress or decompress well it may be an anomaly. Autoencoders are essentially mini neural networks with encoder and decoders.