

# 1 Autoencoder

## 1.1 What is an autoencoder

The special kind of neural networks used in this thesis is called Autoencoder. Its purpose is to reconstruct the content of a given input after said input has been reduced in its dimensionality.

An autoencoder is usually comprised of two main parts: The encoding part or encoder and the decoding part or decoder. The encoder incorporates at least two layers: the input layer and one or more hidden layers. Since autoencoders are usually symmetric in nature, the decoder has the same amount of layers as the encoder. It incorporates the hidden layer(s) and the output layer (see Figure 1).

The idea is that the encoder subsequently *reduces* the dimensionality of the input through several different layers to a compressed representation of the input. Following that the decoder *increases* the dimensionality of the now compressed input until the input and the output have the same dimension. Given that procedure, an autoencoder is classified as an unsupervised learning model.

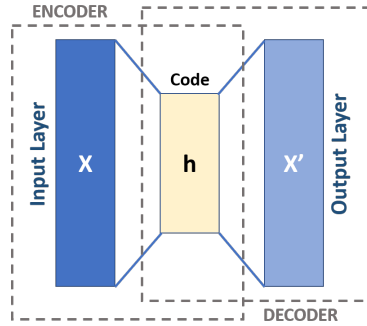


Figure 1: Simple Autoencoder: Platzhalter

## 1.2 Architecture of a simple autoencoder

The two parts of an autoencoder can be mathematically represented as follows:

- The encoder is a function  $f : X \mapsto H$  that maps the  $n$ -dimensional input array  $X \in \mathbb{R}^n$  to the compressed representation  $H \in \mathbb{R}^m$  with  $m$  dimensions, such that  $m < n$ .
- The decoder is a function  $g : H \mapsto Y$  that maps the  $m$ -dimensional compressed representation  $H \in \mathbb{R}^m$  to  $Y \in \mathbb{R}^n$

The autoencoder itself can be represented as the concatenation of these two functions:

- Choose  $f$  and  $g$ , such that  $f, g = \operatorname{argmin}_{f, g} ||X - g(f(X))||$ .

An autoencoder often only has just one hidden layer, but it does not have to be limited to it. More layers can bring certain advantages in some situations like reducing the computational costs<sup>1</sup> or decrease the amount of training data<sup>2</sup>. In this thesis an autoencoder with three hidden layers is used (see Figure 2). The specific architecture of autoencoder is discussed in chapter ...

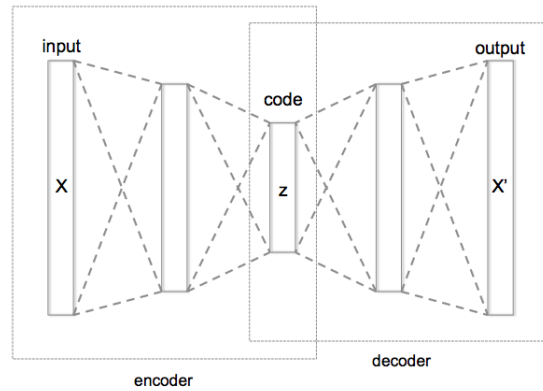


Figure 2: Simple Autoencoder: Platzhalter

---

<sup>1</sup>Wikipedia:Platzhalter

<sup>2</sup>Wikipedia:Platzhalter