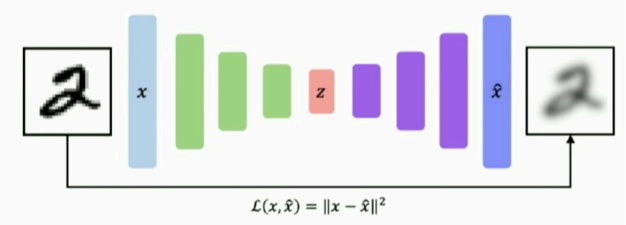
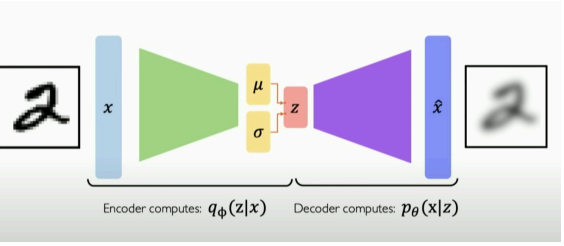
Deep Generative modeling

1. Definition: is a type of ML model capable of not only identifying the patterns but also generating **new** data samples that are similar to the training data. It's called *deep* because it usually involves multiple layers of NN.
2. Aim: is to learn the underlying probability distribution of the training data, so that then it can be able to generate new data.
    a. Latent variable model: this kind of model is used to learn the underlying structure of the data by modeling the probability distribution of the observed variables given the latent variables.
    b. Autoencoders: are compression models that create a feature representation of data through a bottlenecked compressed hidden layer. It also can be used for generative tasks, as the decoder part can be employed to generate new data points by sampling from the learned compressed representation.

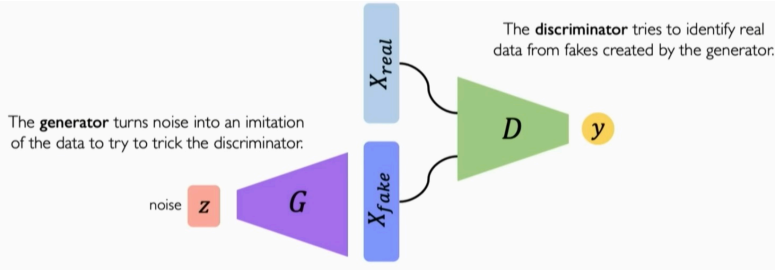


3. Examples:
    a. **Variational Autoencoders (VAEs):** These types of models use latent variable models to generate new data samples, they encode the input data into a latent variable space and decode the latent variable back into the original data space to generate new data.
    Priors on the latent distribution and **reparameterizing the sampling layer** are techniques used in this kind of model to achieve continuity and completeness in the encoded latent space.



    b. **Generative Adversarial Networks (GANs)**

The objective is to make fake samples from scratch, for example: a face of a person that doesn't exist. To do this we take noise as an input and train the model to the data distribution.
We essentially train 2 NN at the same time. A generator and a discriminator that compete against each other.



Feedback: Discriminator feedback is used to update the weights of the generator and discriminator. The generator adjusts its weights to generate data that is more difficult for the discriminator to distinguish, and the discriminator adjusts its weights to improve its discriminating ability. The global optimum is found when the generator generates fake data so well that it is difficult to tell the discriminator that they are not real.

Conditional GANs: work similarly but instead of generating data from white noise, they generate it from a data manifold X. Takes additional input, commonly tags or any other relevant conditional information. This allows specific data to be generated based on the information provided.

CycleGANs: used to perform translations between two data domains without requiring matching data pairs and employ cyclic consistency.