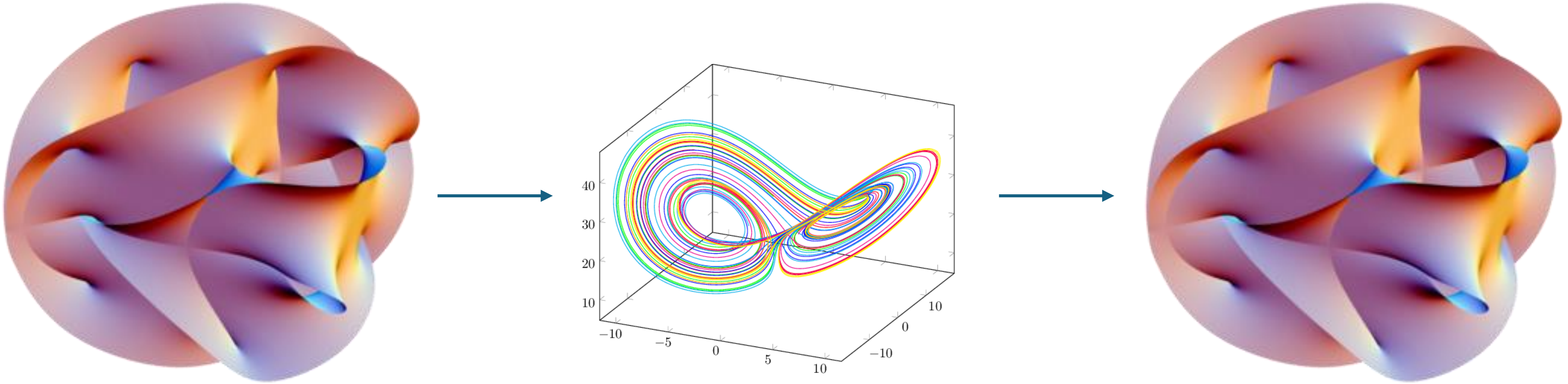
ML Enabled Representation and Visualization



Dr. Bryan Scott

LSST Discovery Alliance Data Science Fellowship Program Session 22

(partially) adapted from a tutorial on VAEs by Charles Kenneth Fisher and Raghav Kansal

Overview

Representations in Machine Learning:

Why representation learning? What is the connection between representation learning and visualization?

Specific Example: Autoencoders

Overview of autoencoder architectures and latent space representations

Revenge of the Bayesians: Variational Autoencoders (VAEs)

How can we make this probabilistic? Why would we want to?

Overview

Representations in Machine Learning:

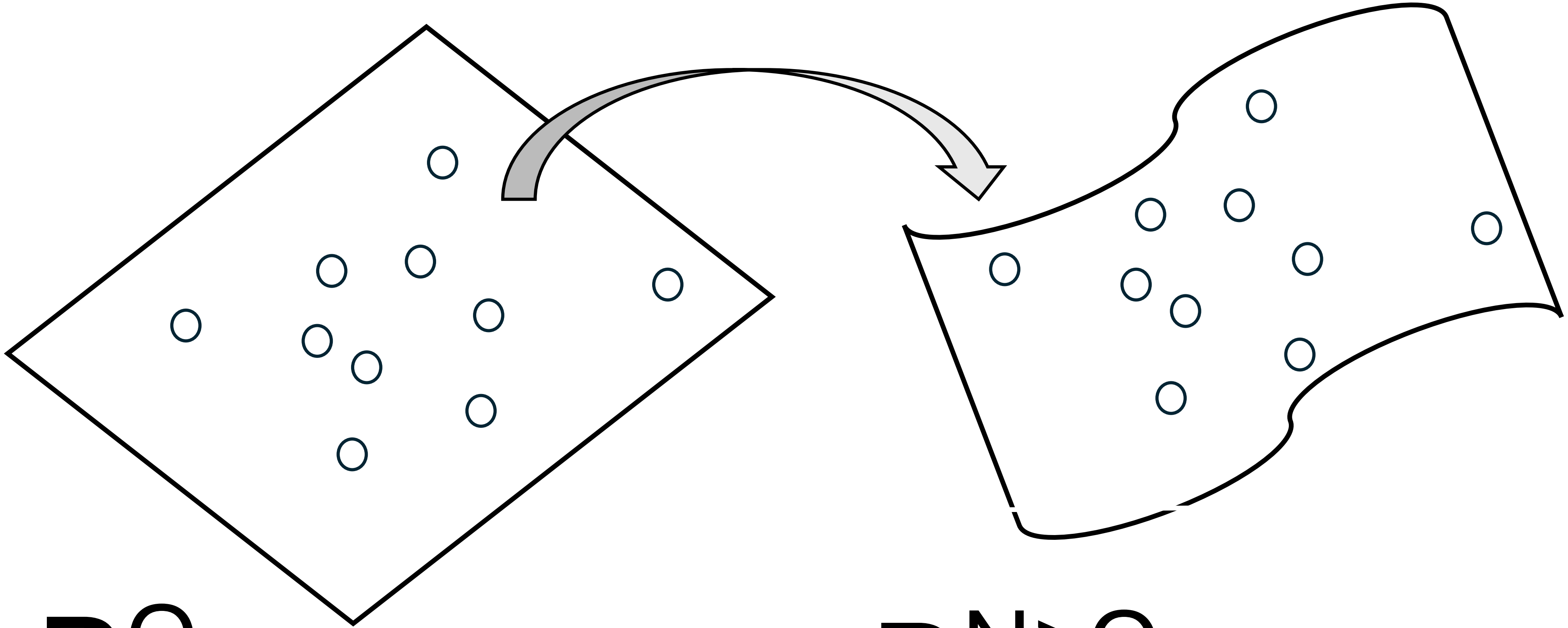
Why representation learning? What is the connection between representation learning and visualization?

Specific Example: Autoencoders

Overview of autoencoder architectures and latent space representations

Revenge of the Bayesians: Variational Autoencoders (VAEs)

How can we make this probabilistic? Why would we want to?



R^Q

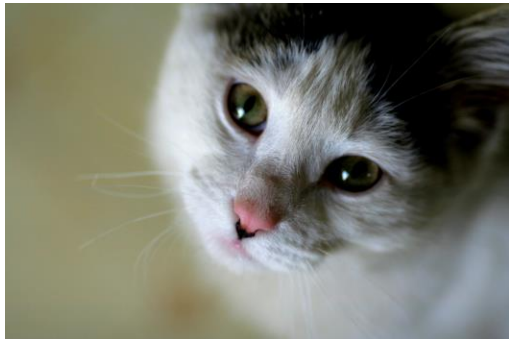
Data is “generated” in this space.

$R^{N>Q}$

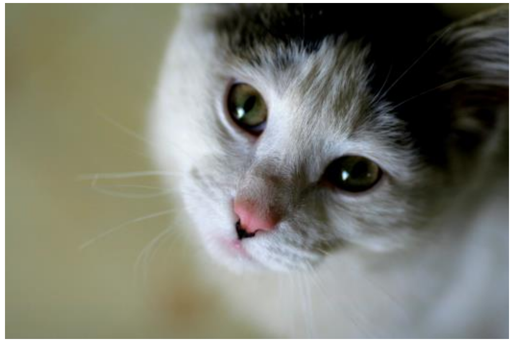
Observations are made in this space.

Supervised Learning

Unsupervised Learning

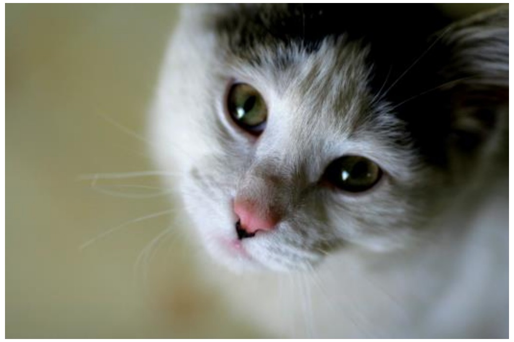


Unsupervised Learning



Representation Learning, a secret third thing?



Representation Learning, a secret third thing?

Representation Learning and Visualization

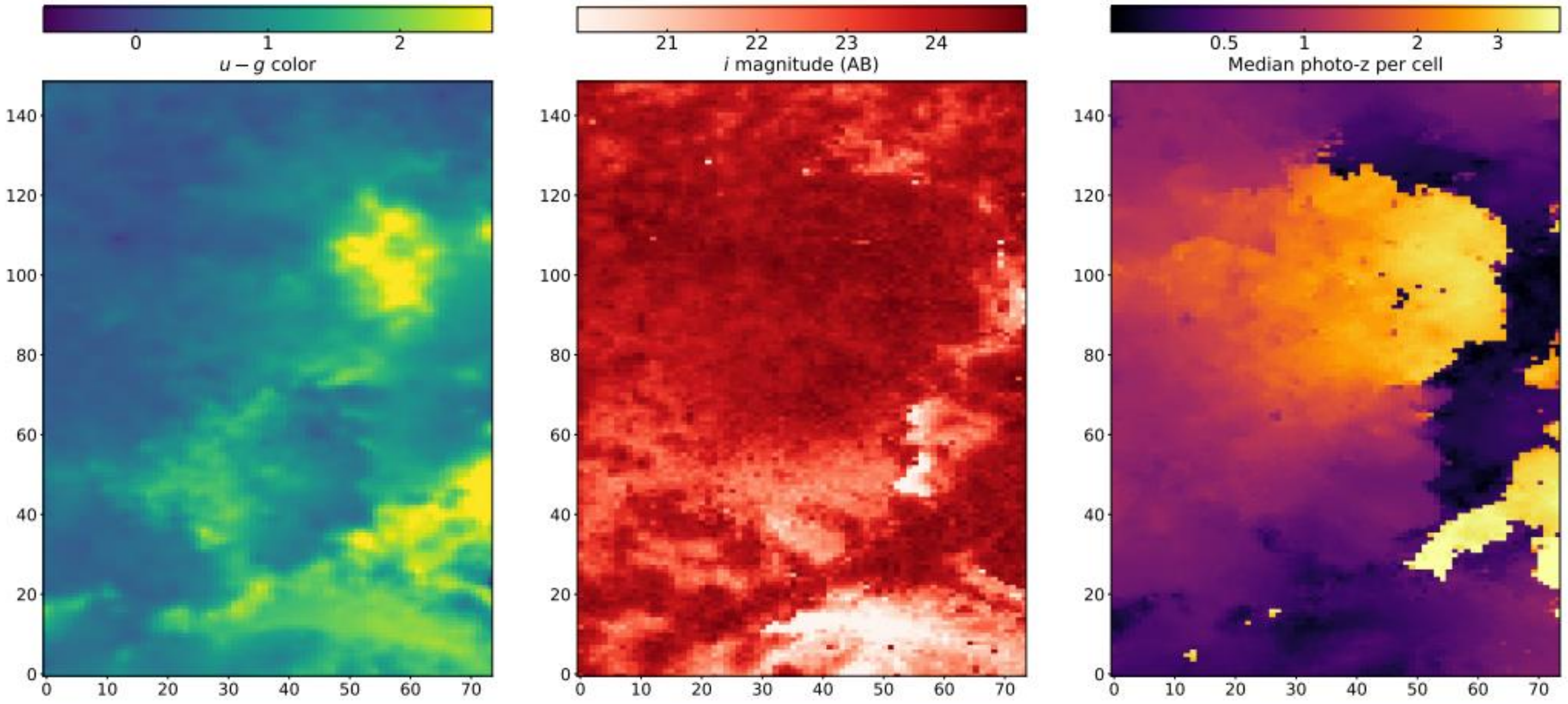


Figure 4. Illustration of the updated self-organized map (SOM) used for the C3R2 DR2 source selection and analysis. Each cell of the SOM represents a particular spectral energy distribution (SED) that shows up with regularity in the deep field data. The axes should be thought of as indices to parts of the high-dimensional galaxy color space. *Left:* The map colored according to one of the features it encodes, namely the u-g color. *Middle:* The map colored according to the median i band magnitude of galaxies occupying each cell. It is clear that the typical magnitude is strongly color dependent. *Right:* The SOM colored by the median photometric redshift of galaxies occupying each cell. Note the topological nature of the SOM: similar SEDs group together, producing relatively smooth variation of photo-z with position on the map.

Masters et al., 2019

Overview

Representations in Machine Learning:

Why representation learning? What is the connection between representation learning and visualization?

Specific Example: Autoencoders

Overview of autoencoder architectures and latent space representations

Revenge of the Bayesians: Variational Autoencoders (VAEs)

How can we make this probabilistic? Why would we want to?

What is an autoencoder?

An autoencoder is a type of algorithm with the primary purpose of learning an "informative" ***representation*** *of the data that can be used for different applications a by learning to* ***reconstruct*** *a set of input observations well enough*.

The **reconstruction** is the “*auto*” part

The **representation** is the “*encoding*” part.

Autoencoders are an ‘architecture’ for performing the representation learning task. The specific implementation can take many forms.

How to build an autoencoder?

Autoencoders are approximations to the identity function.

How to build an autoencoder?

Autoencoders are approximations to the identity function.

Given a dataset x, what function x' = f(x) yields x = x'? Neural networks are universal function approximators, so a neural network should be a good choice for approximating the simplest possible function – the identity.

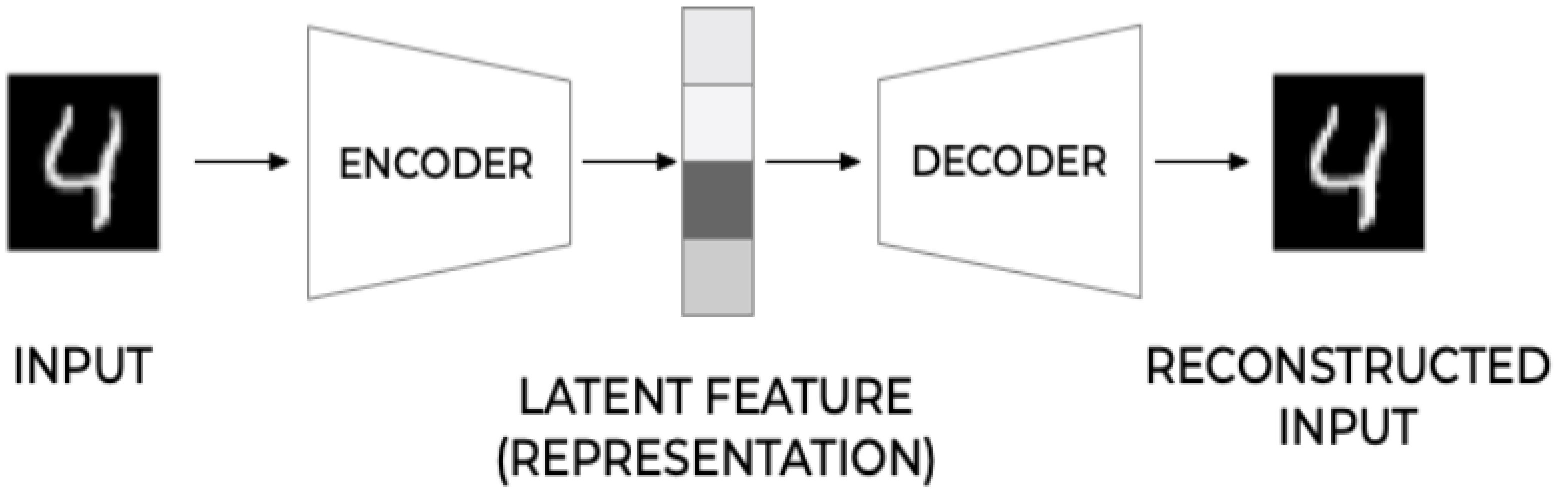
How to build an autoencoder?

Autoencoders are approximations to the identity function.

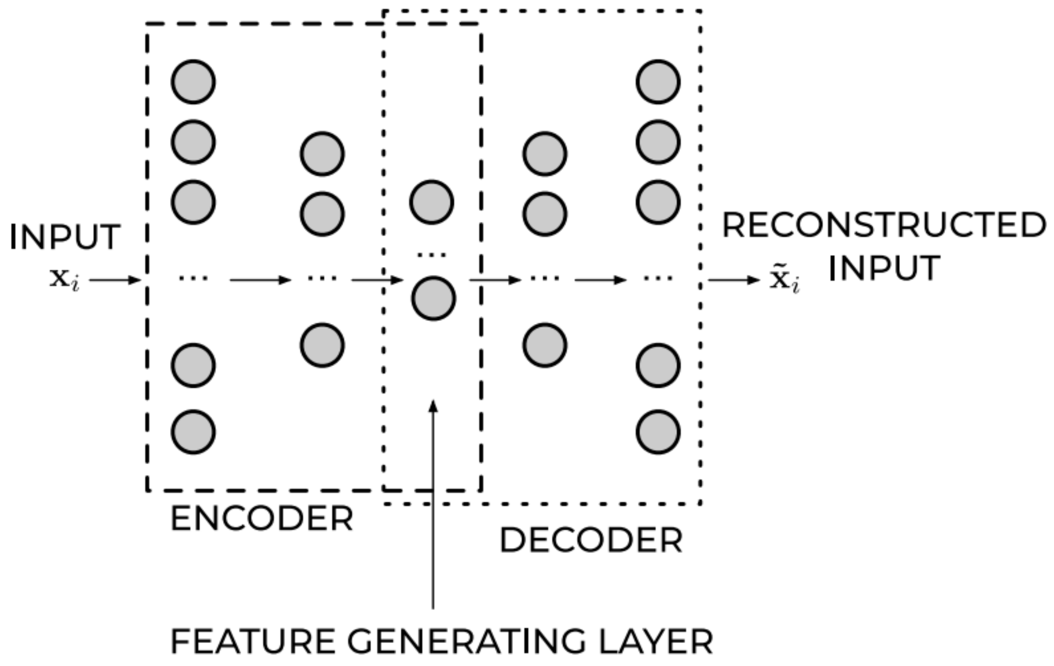
Given a dataset x, what function x’ = f(x) yields x = x’? Neural networks are universal function approximators, so a neural network should be a good choice for approximating the simplest possible function – the identity.

By choosing a neural network architecture, we will introduce additional degrees of freedom that can yield *representative* or *interpretable* **compressions** of the data.

Schematic Representation



Dimensionality Reduction in Autoencoders:



The encoder can reduce the number of dimensions of the input observation (n) and create a learned representation (hi)of the input that has a smaller dimension q < n. This learned representation is enough for the decoder to reconstruct the input accurately (if the autoencoder training was successful as intended).

Overview

Representations in Machine Learning:

Why representation learning? What is the connection between representation learning and visualization?

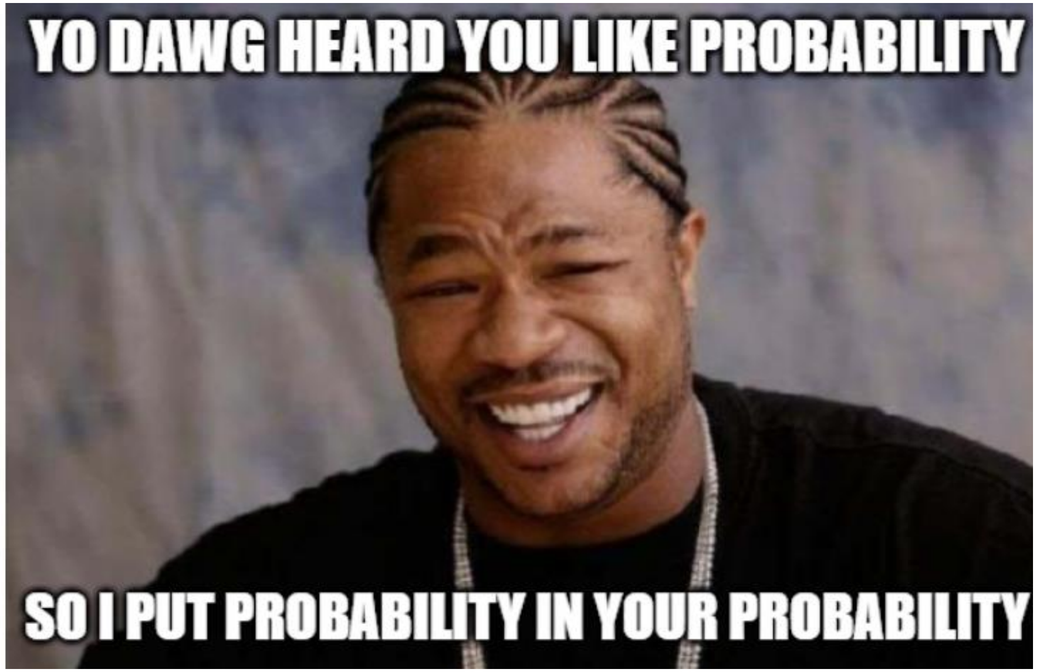
Specific Example: Autoencoders

Overview of autoencoder architectures and latent space representations

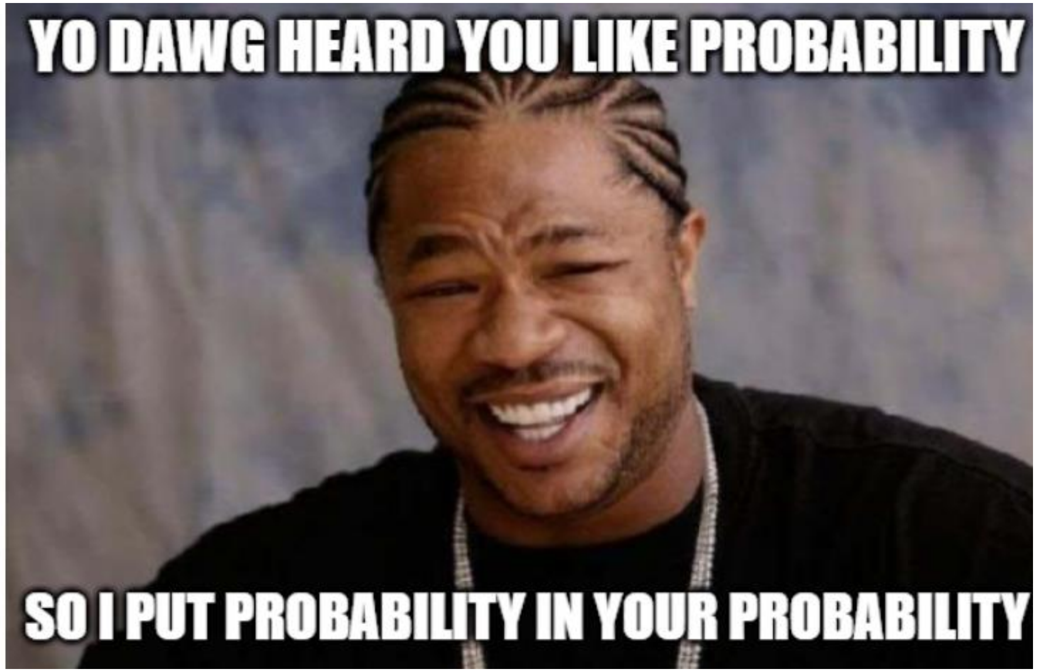
Revenge of the Bayesians: Variational Autoencoders (VAEs)

How can we make this probabilistic? Why would we want to?

Revenge of the Bayesians: Variational Autoencoders



Revenge of the Bayesians: Variational Autoencoders

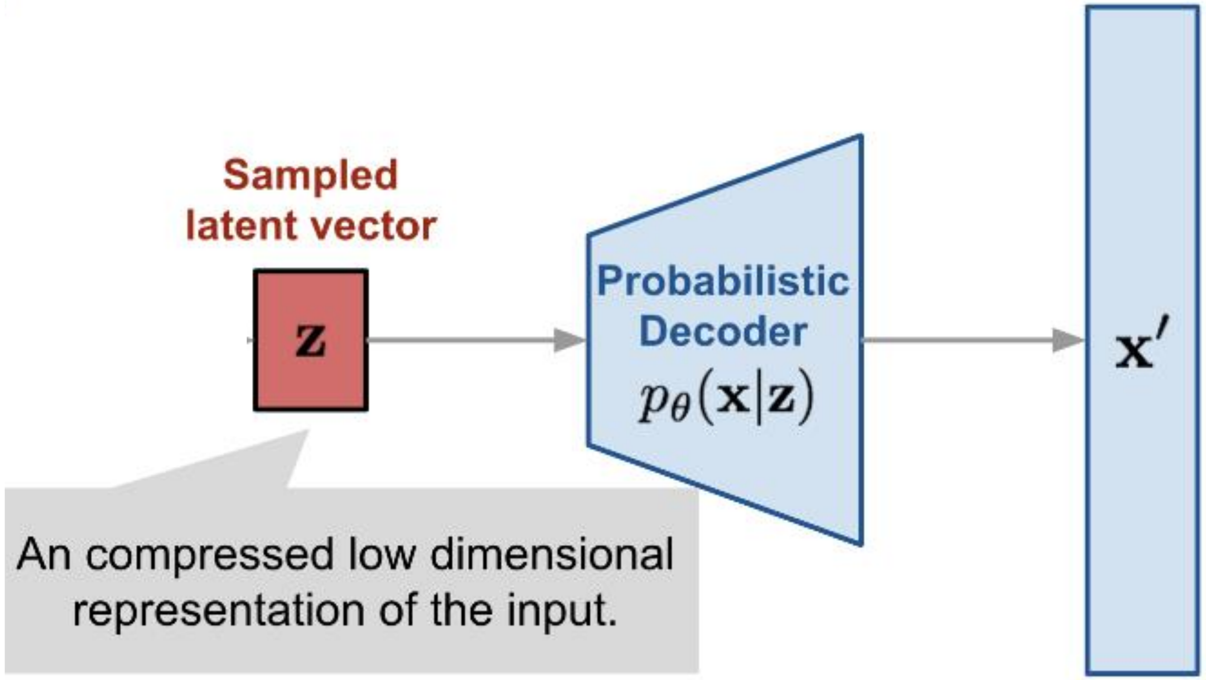


Goal of a VAE: Learn the Joint Distribution

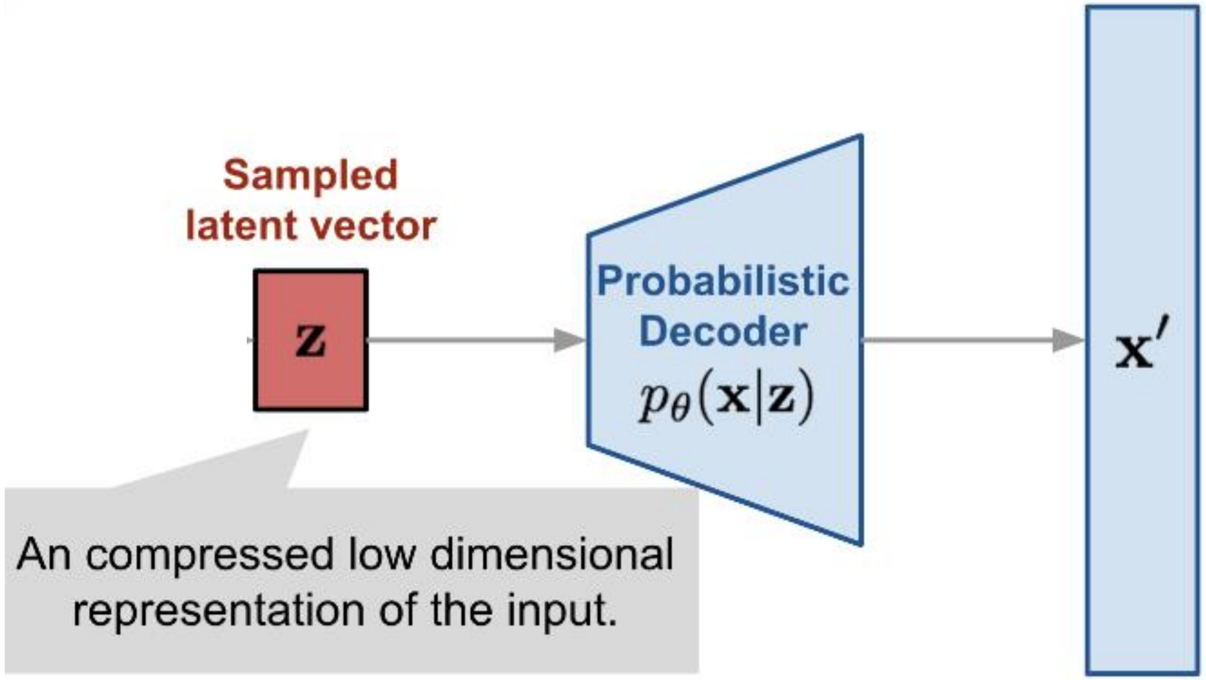
Estimate

$$p_\theta(\mathbf{x}, \mathbf{z}) = p_\theta(\mathbf{x}|\mathbf{z})p(\mathbf{z}).$$

VAE Latent Space and Decoder



VAE Latent Space and Decoder



"Generative Model"

Goal of a VAE: Learn the Joint Distribution

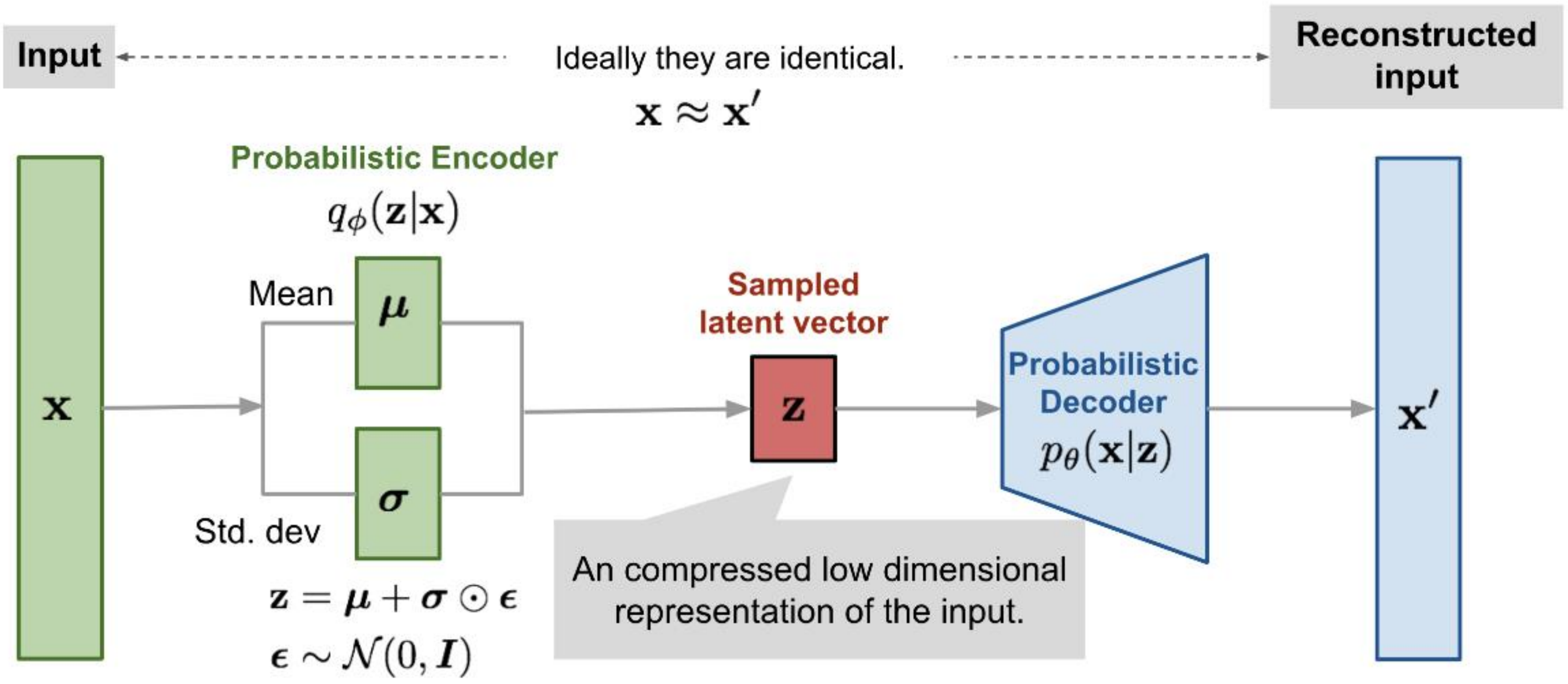
Estimate

$$p_\theta(\mathbf{x}, \mathbf{z}) = p_\theta(\mathbf{x}|\mathbf{z})p(\mathbf{z}).$$

Assuming,

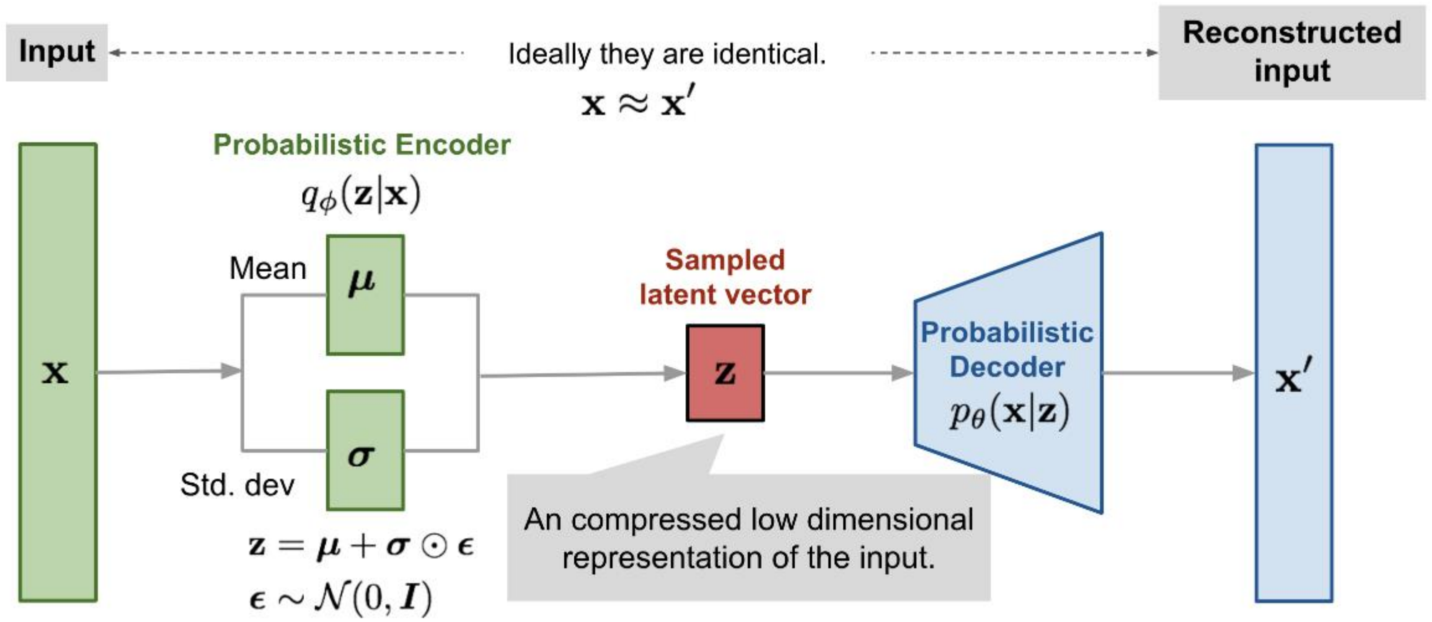
$$p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

VAE Encoder



Input
Ideally they are identical.
$\mathbf{x} \approx \mathbf{x}'$
Reconstructed input
Probabilistic Encoder
$q_\phi(\mathbf{z}|\mathbf{x})$
$\mathbf{x}$
Mean
$\boldsymbol{\mu}$
Std. dev
$\boldsymbol{\sigma}$
$\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \boldsymbol{\epsilon}$
$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{I})$
Sampled latent vector
$\mathbf{z}$
An compressed low dimensional representation of the input.
Probabilistic Decoder
$p_\theta(\mathbf{x}|\mathbf{z})$
$\mathbf{x}'$

Loss Function Terms



$$-D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})|p(\mathbf{z})) = \frac{1}{2}\sum_{j=1}^{J}\left(1 + \log \sigma_j^2(\mathbf{x}) - \mu_j^2(\mathbf{x}) - \sigma_j^2(\mathbf{x})\right.$$

What have we done? Variational Posterior Estimation (A preview of Session 23)

Dealing with Intractable Posteriors: Evidence Lower Bound

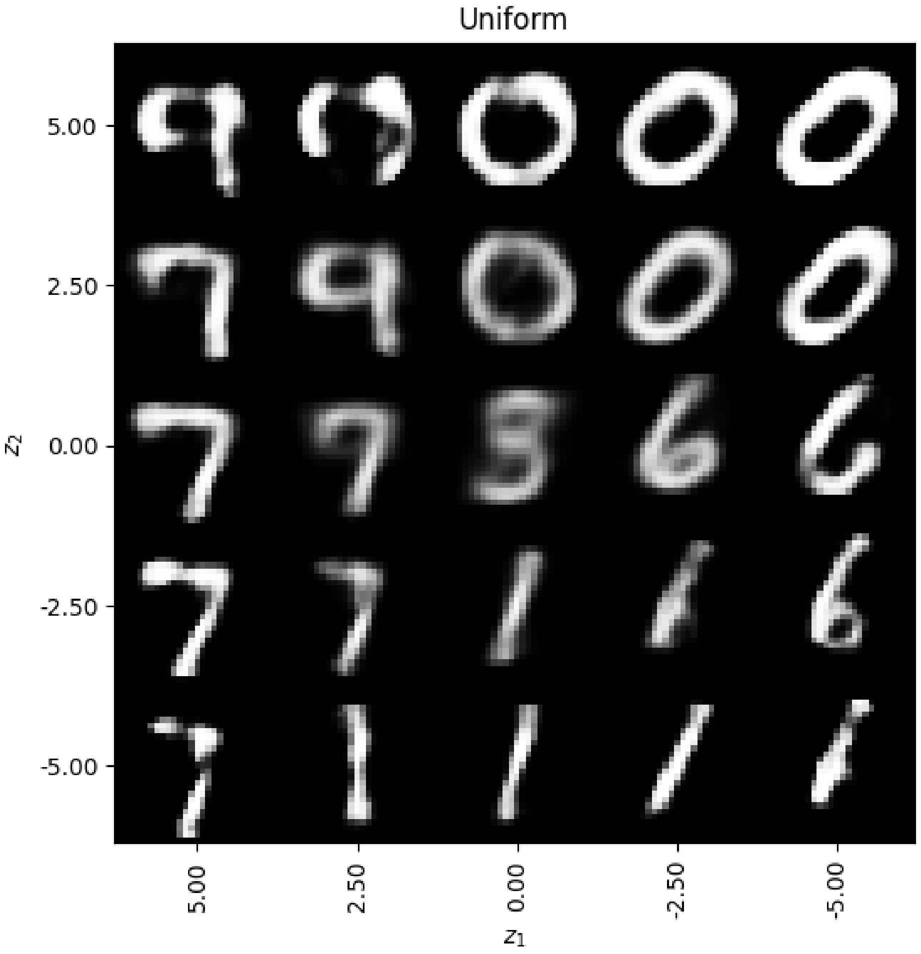
$$\begin{aligned}\mathrm{ELBO}(q) &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})] \\ &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})] - \mathrm{KL}[q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})].\end{aligned}$$

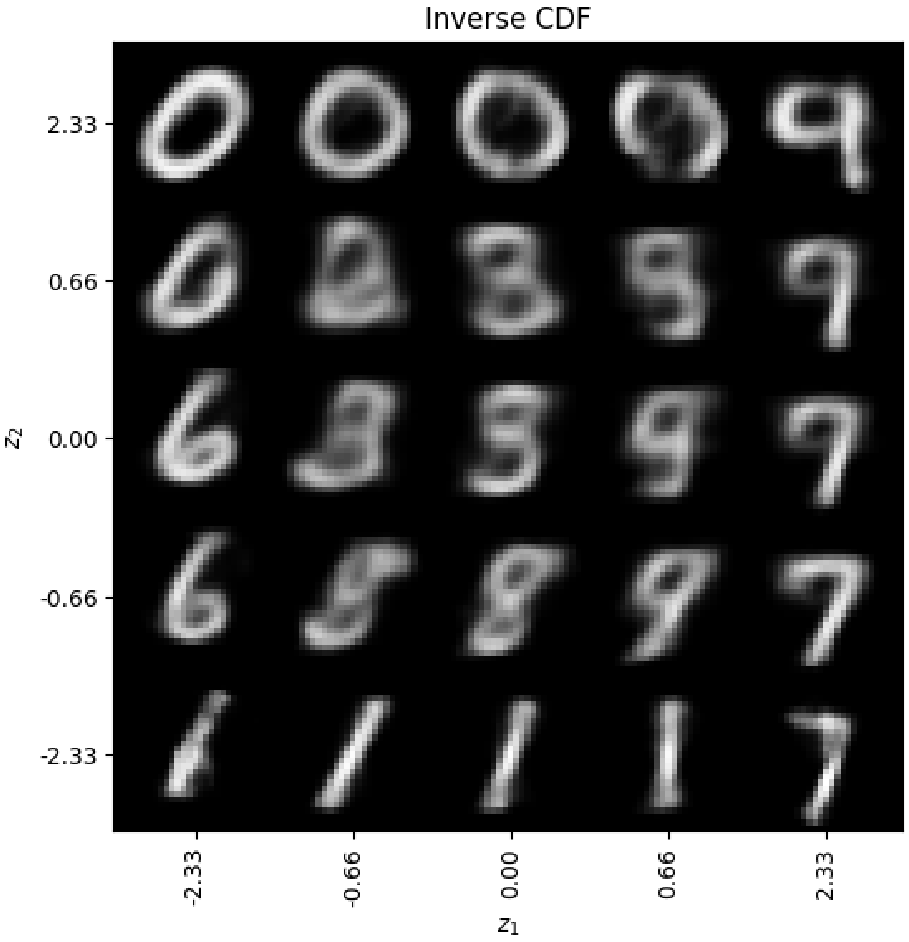
"Reparameterization Trick"

$$\begin{aligned}\nabla_\phi \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[f(\mathbf{x},\mathbf{z})] &= \nabla_\phi \mathbb{E}_{p(\epsilon)}[f(\mathbf{x}, g_\phi(\mathbf{x},\boldsymbol{\epsilon}))] \\ &= \mathbb{E}_{p(\epsilon)}[\nabla_\phi f(\mathbf{x}, g_\phi(\mathbf{x},\boldsymbol{\epsilon}))].\end{aligned}$$

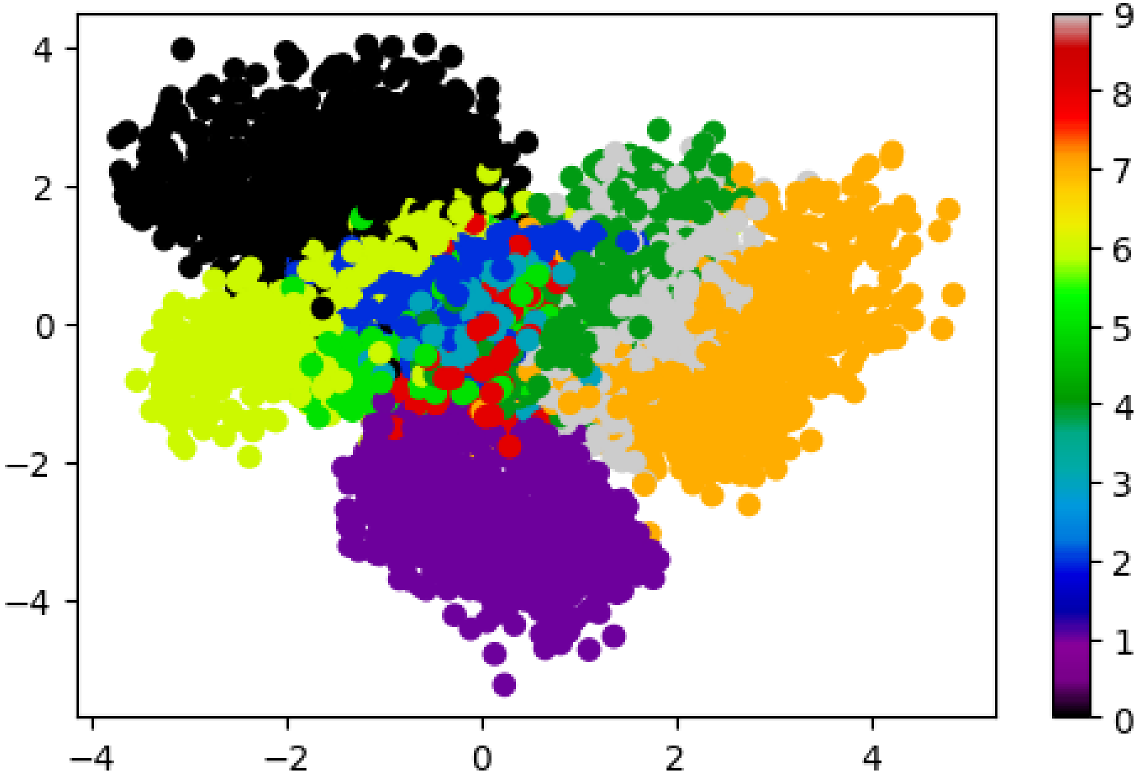
Worked Example: MNIST Dataset

Latent Space Visualizations





Latent Space Visualizations



Problem Notebook

Using a Galaxy Zoo Hubble Space Telescope Image set, train a VAE for HST galaxy images.

Then, generate some examples from the latent space.

Use these to draw some conclusions about the learned representation of the image data. Does the two-dimensional latent space have a clear structure? What do the axes correspond to?