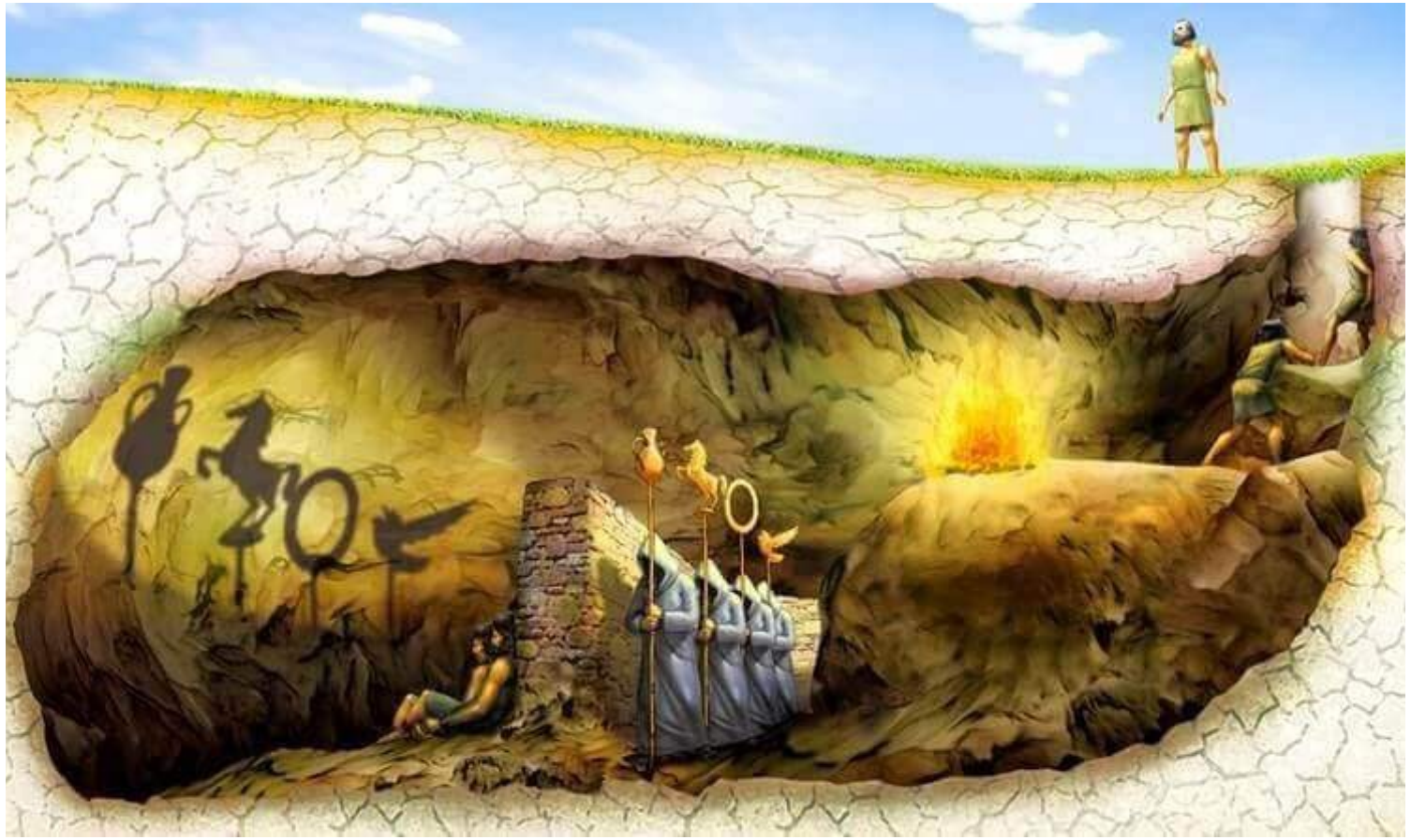
Computer Vision (CS 419/619) Spring 2024

Autoencoders



Myth of the Cave (*Plato's Allegory of the Cave)* [1]

- The aim of generative models is to learn or understand these latent variables, even when we are just given the observed variables/data.

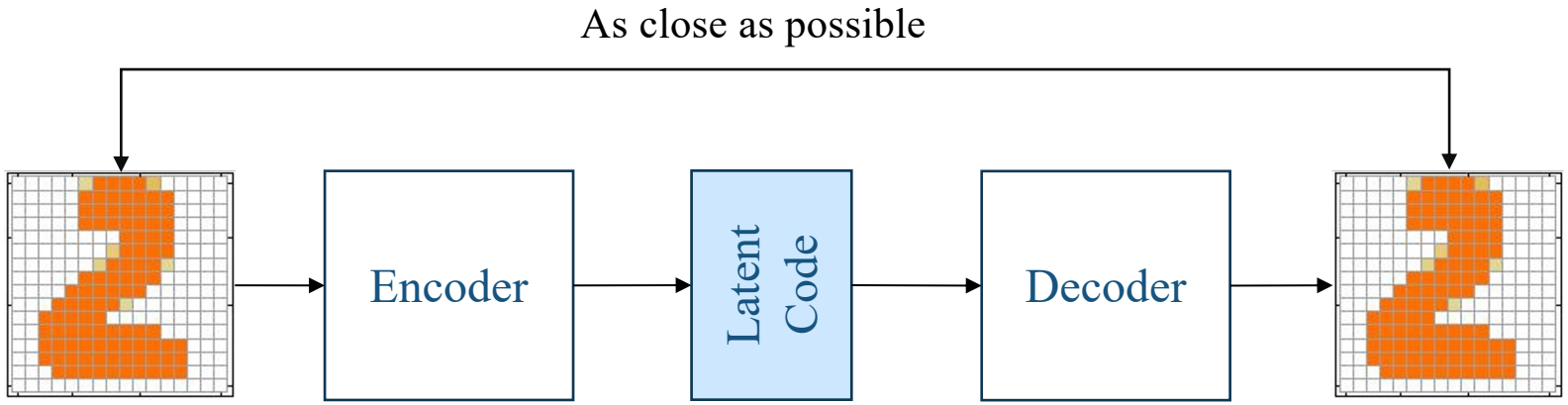
- The story begins with prisoners who have lived their entire lives chained inside a cave.
- Behind the prisoners is a fire, and between the fire and the prisoners are people carrying puppets or other objects.
- These cast shadows on the opposite wall.
- The prisoners watch these shadows, believing this to be their reality as they have known nothing else.

- From the prisoners' perspective, the shadows are the "observed variables".
- These variables can be perceived and measured.

- Behind them are the true objects that are casting the shadow.
- From the prisoners' perspective, these objects are hidden from them.
- They are can termed as "latent variables".

1. https://www.studiobinder.com/blog/platos-allegory-of-the-cave/

Generating new images



Randomly generate a vector as code

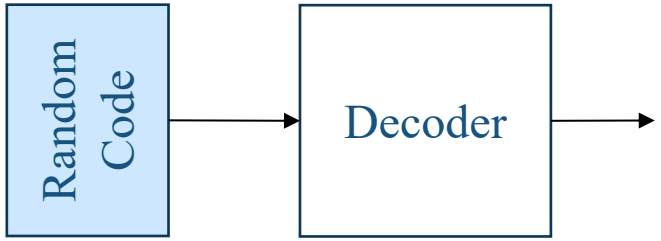


Image ?

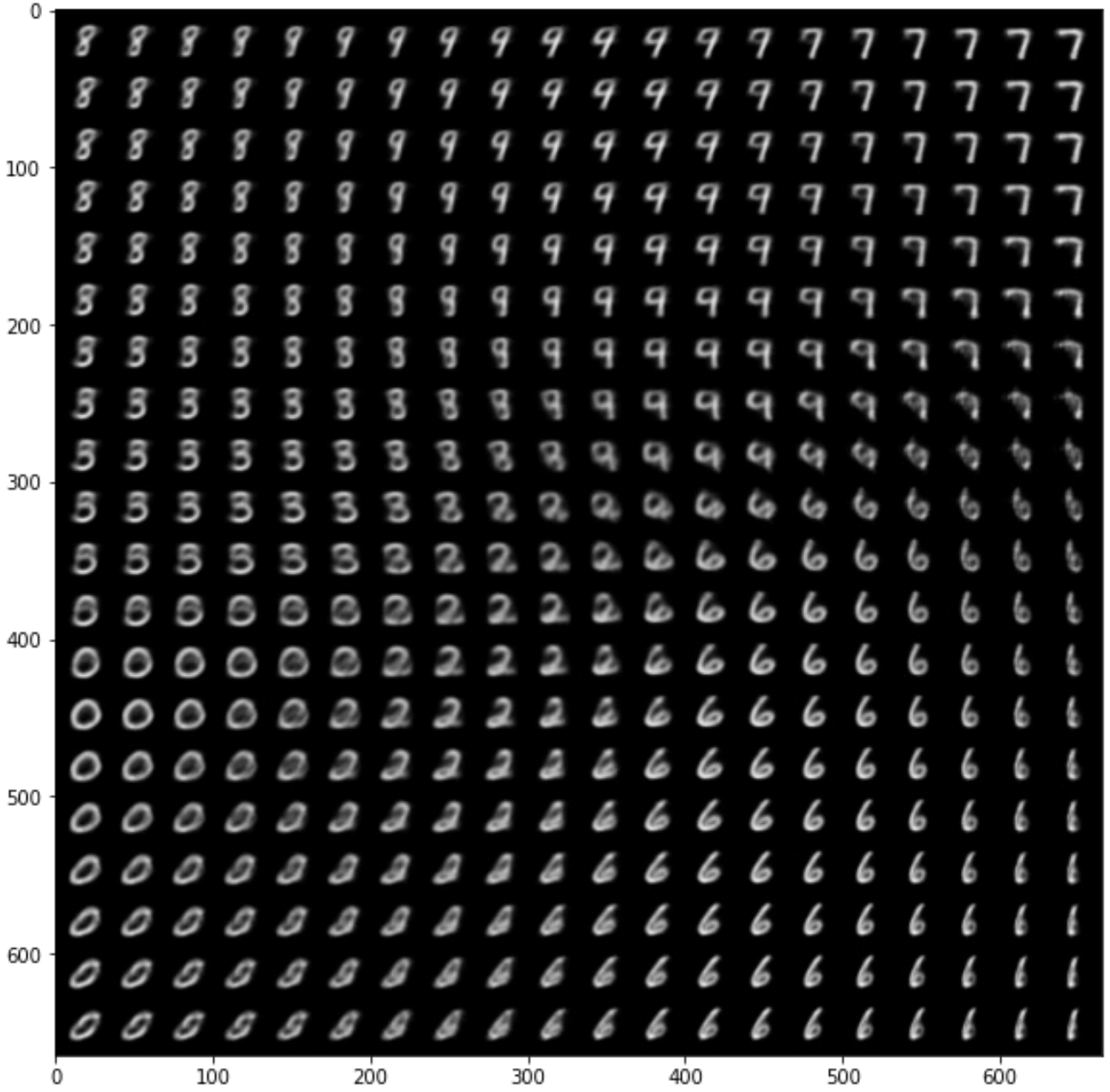
Generating new images

Notice the features along each axis:

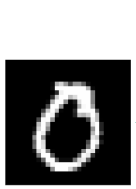
- Along the x-axis if move from 0 to 600.
  - There is a shift in feature
  - The images tend to go from circle to a straight line

- Along the y-axis
  - The above values tend (~600) to show a loop on the top of the digit
  - The above values (~0) tend to show a loop on the bottom of the digit.

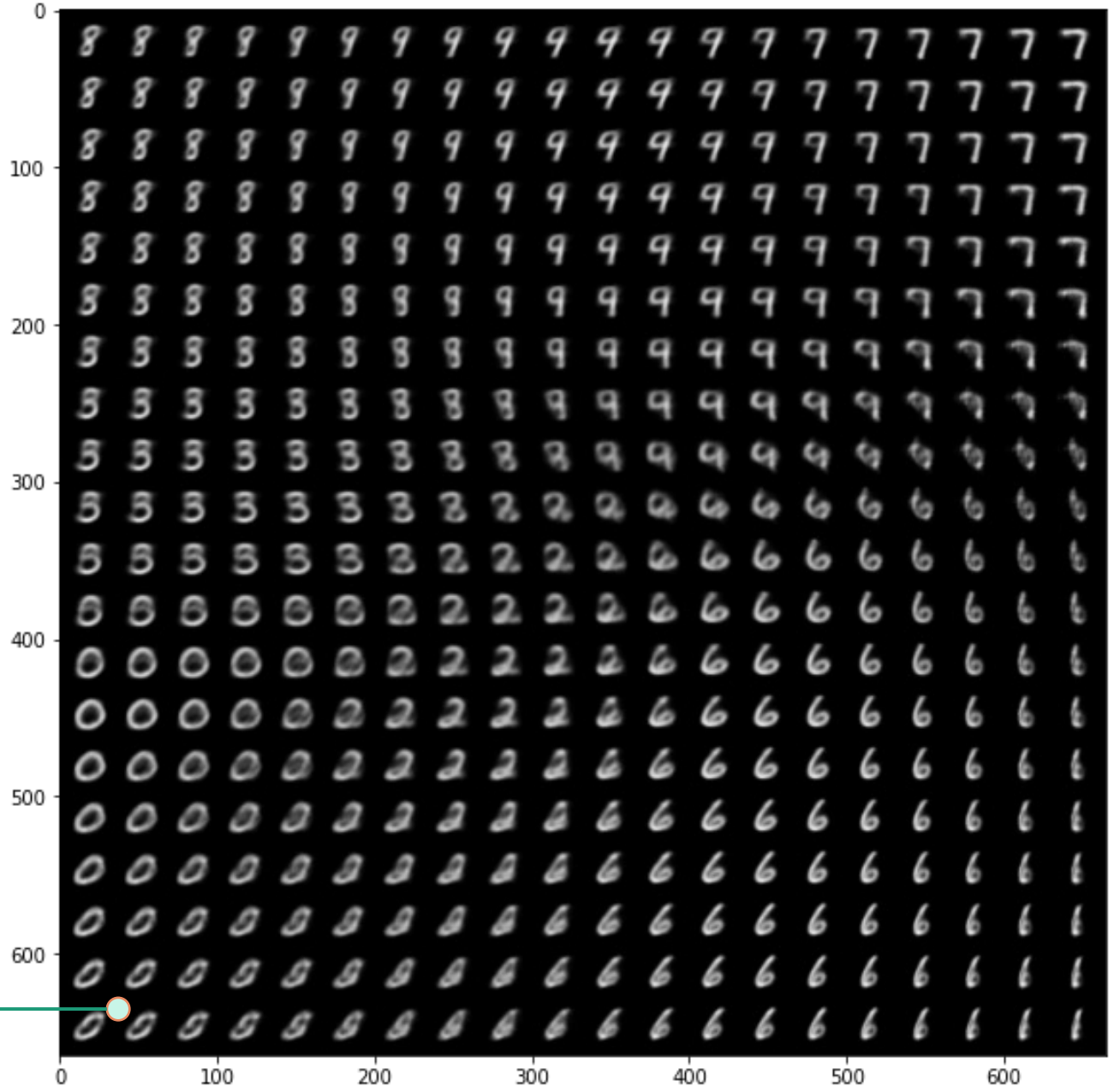
- Grid of values in the latent space.
- The axes denotes the values in the latent vector.
- The images (*coordinates*) denote the reconstructed image from that vector.



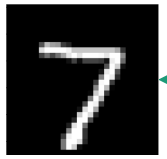
Generating new images



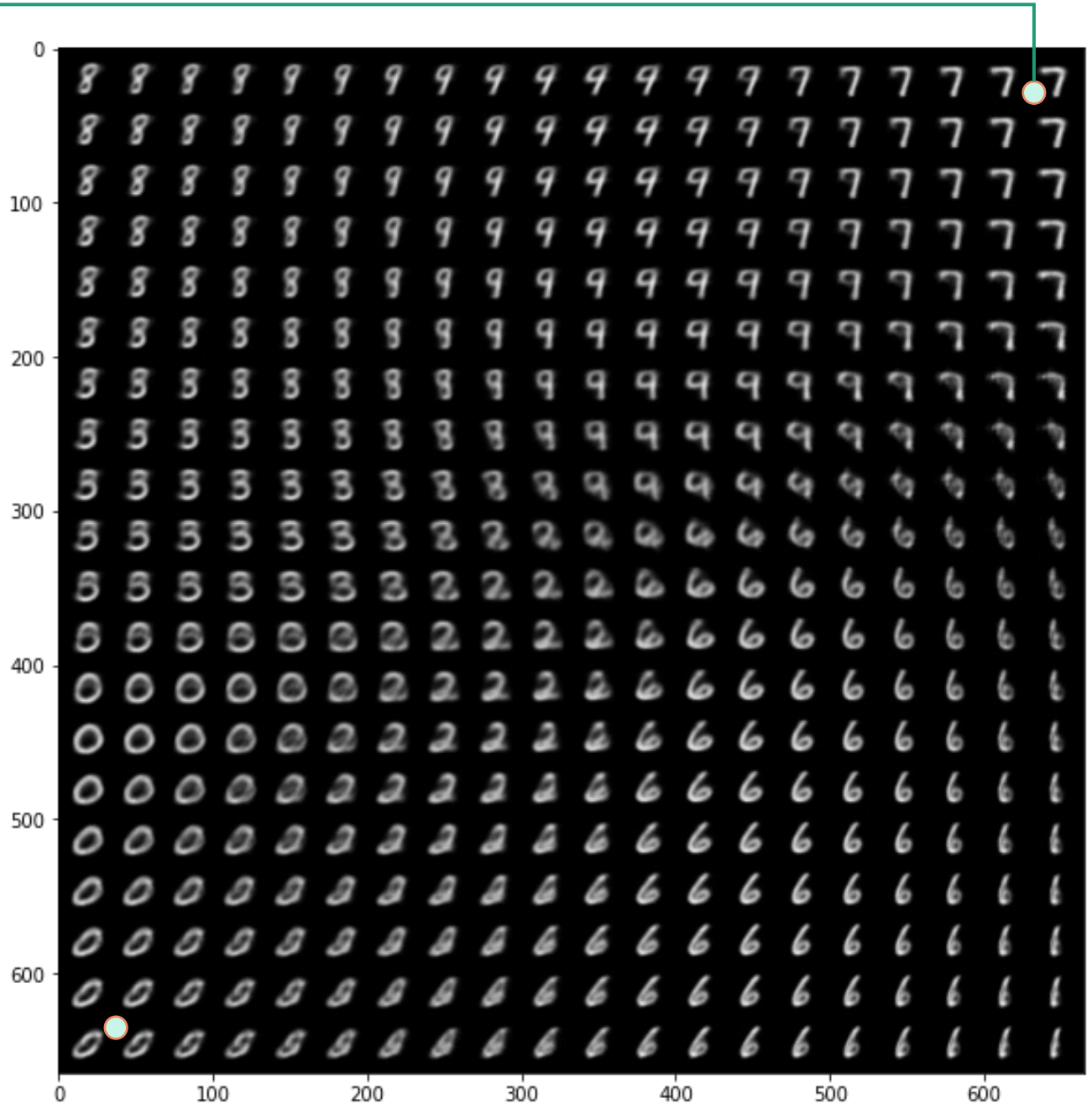
Decoder



Generating new images

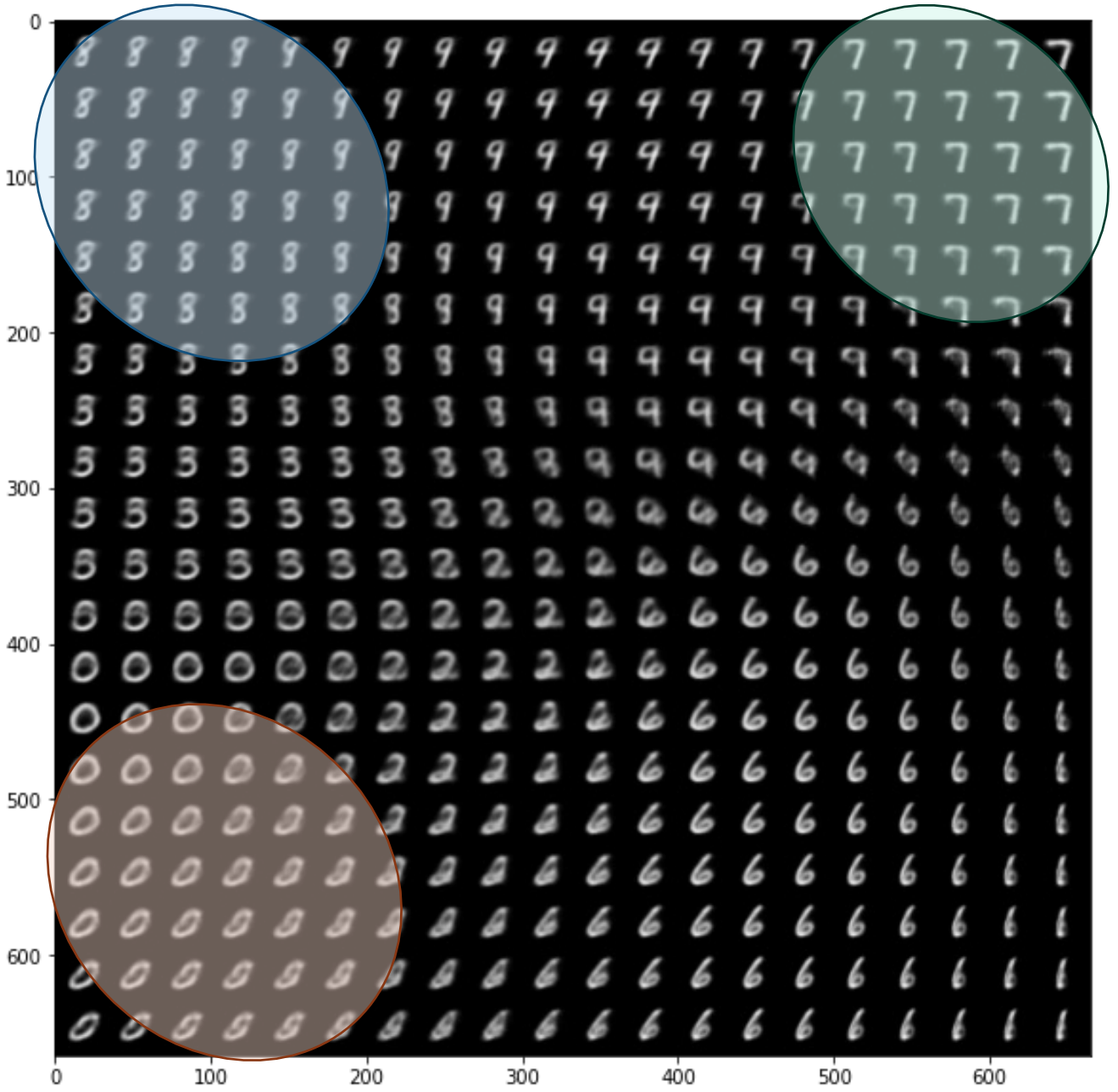


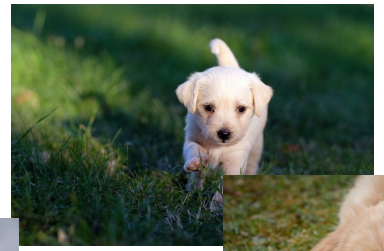
Decoder

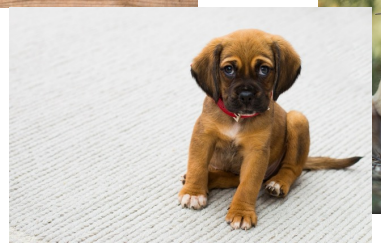
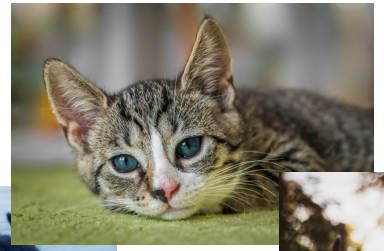


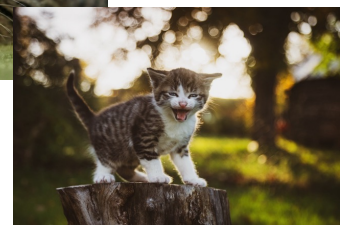
Generating new images

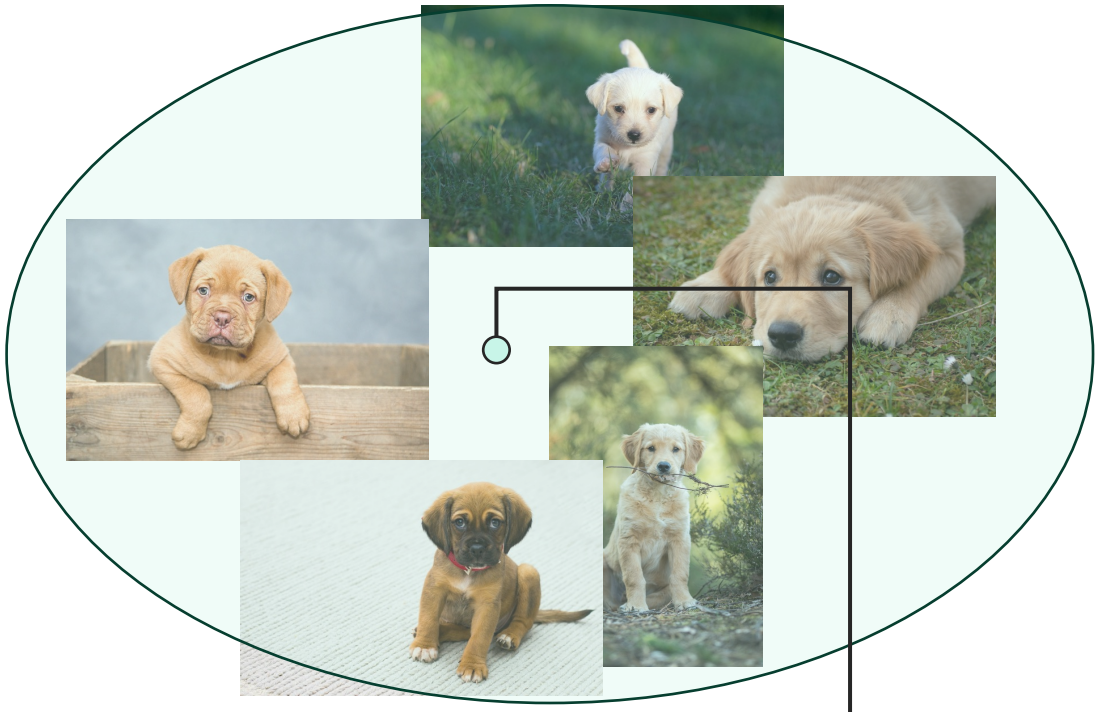
- Autoencoder tries to model the probability density of location of each class.
- Some *'special'* types of autoencoder try to model only the parameters of probability distribution for each class.
- *We will revisit the second point in a bit.*

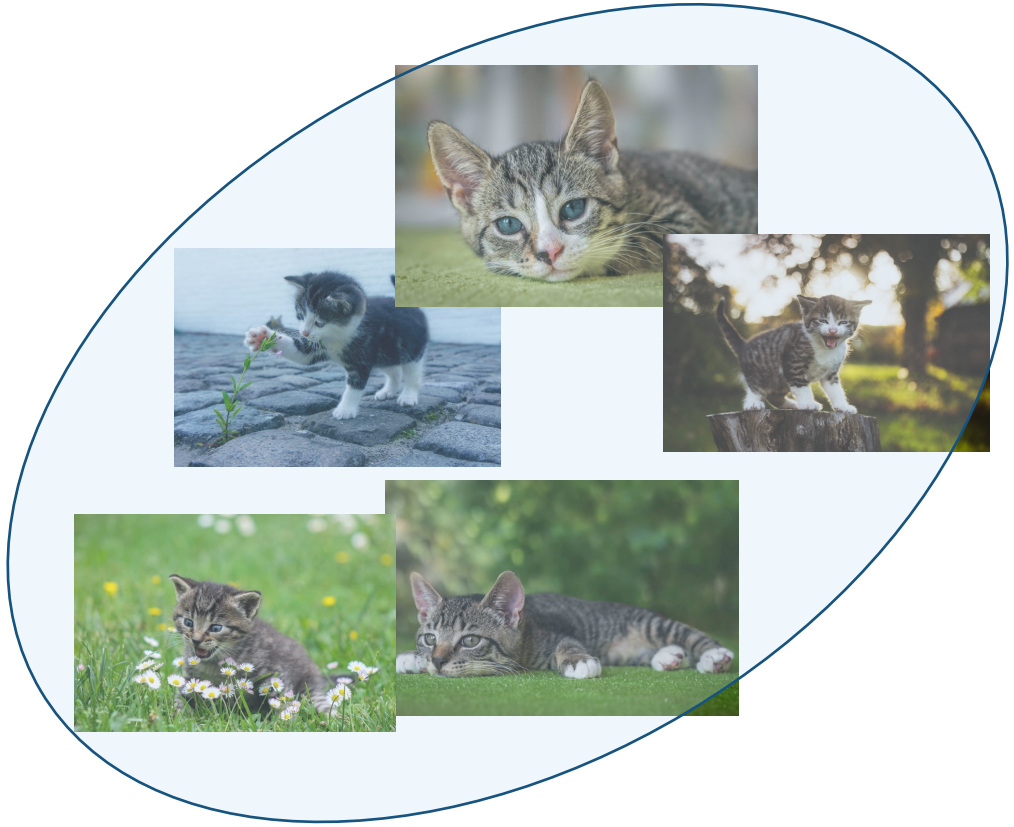




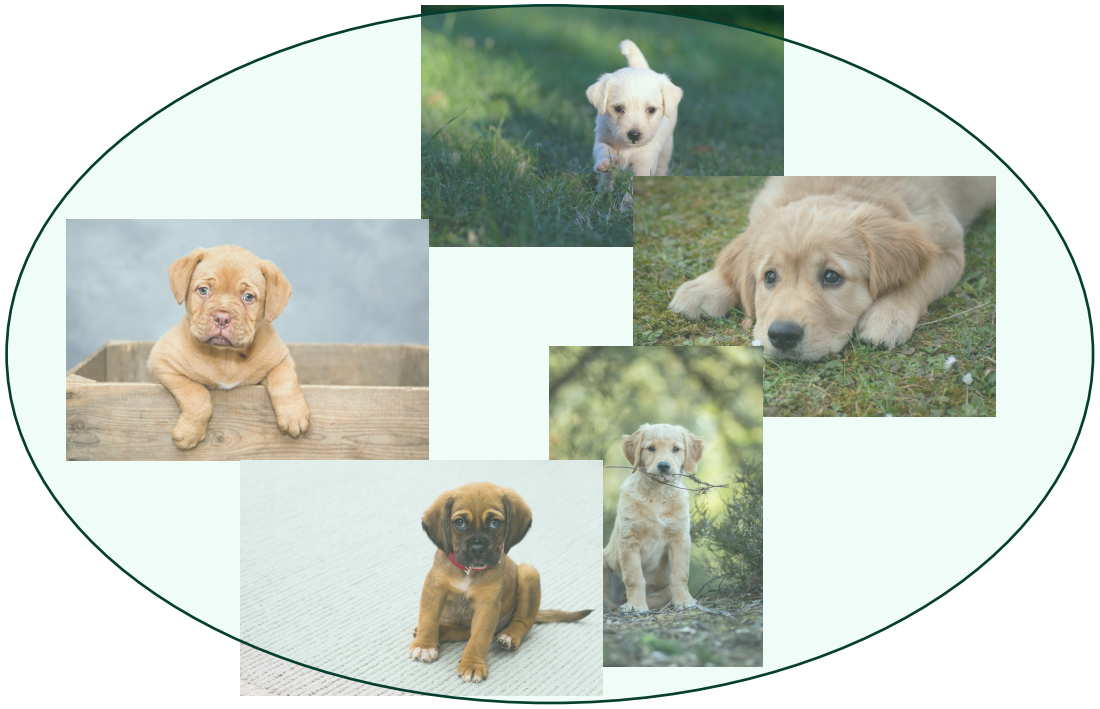



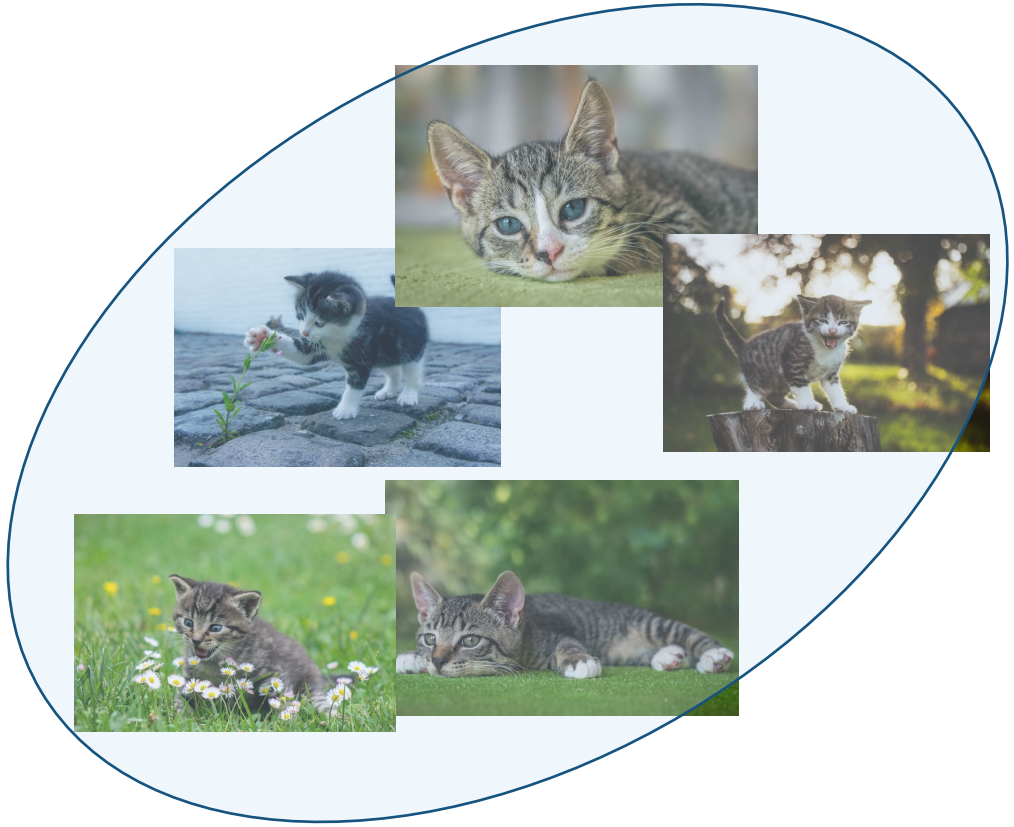


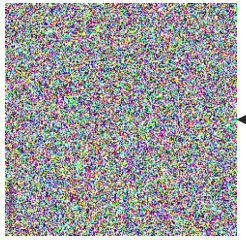




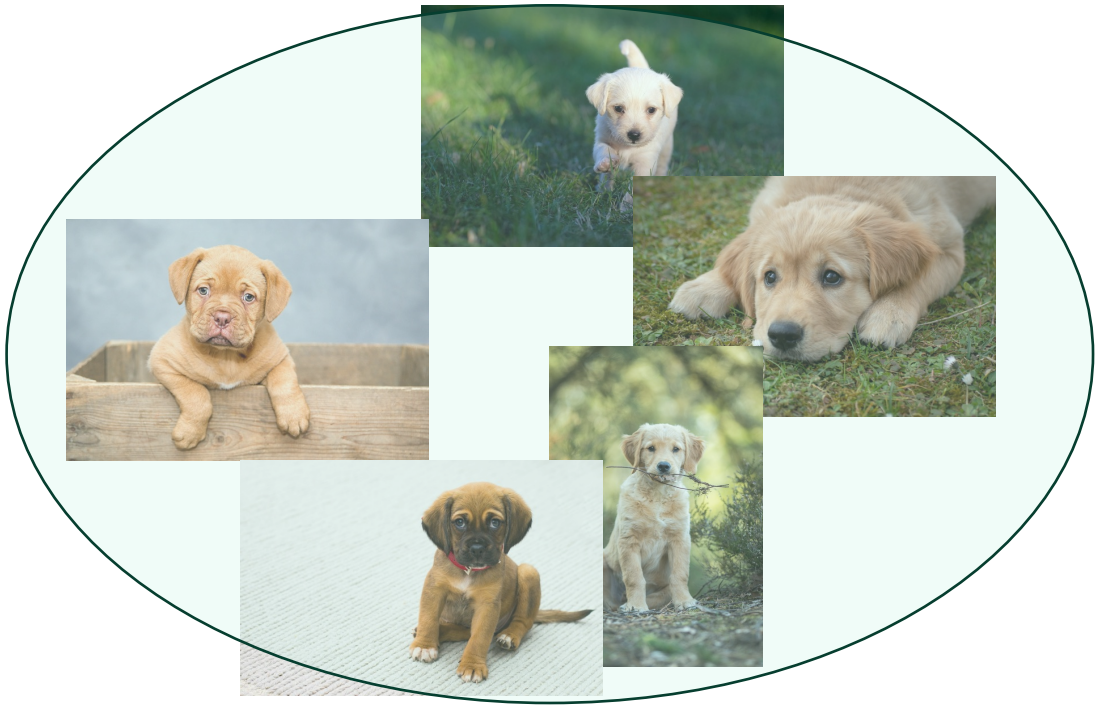
Decoder

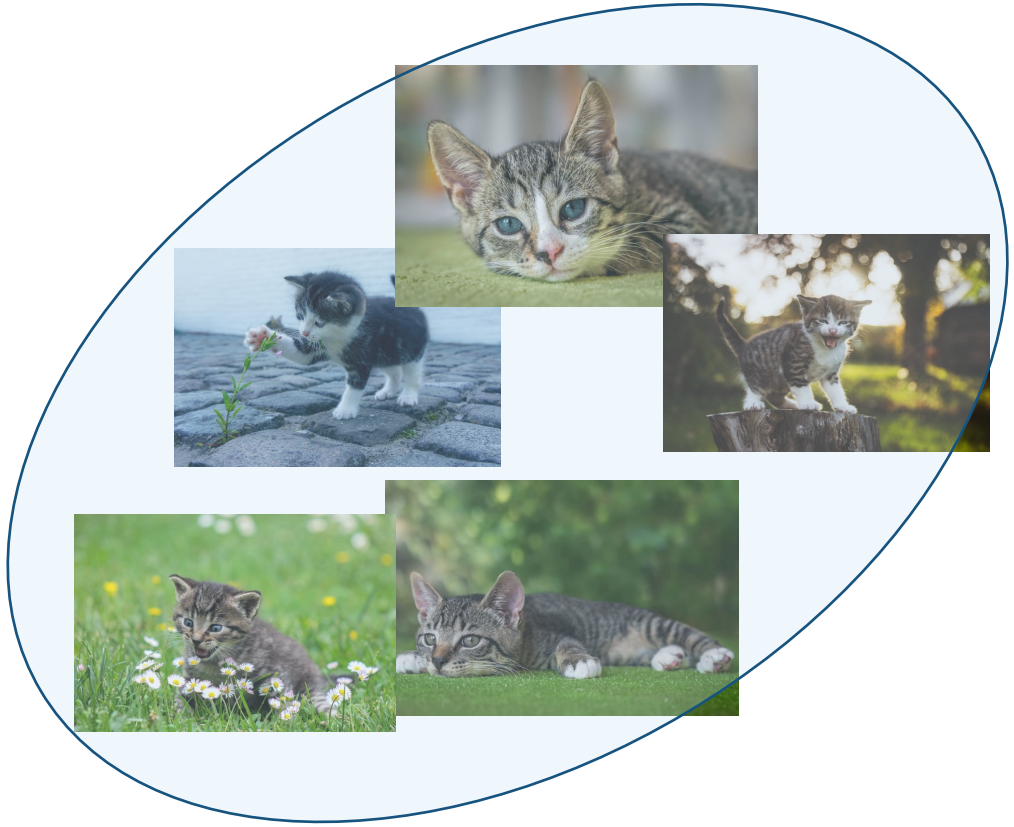






Decoder



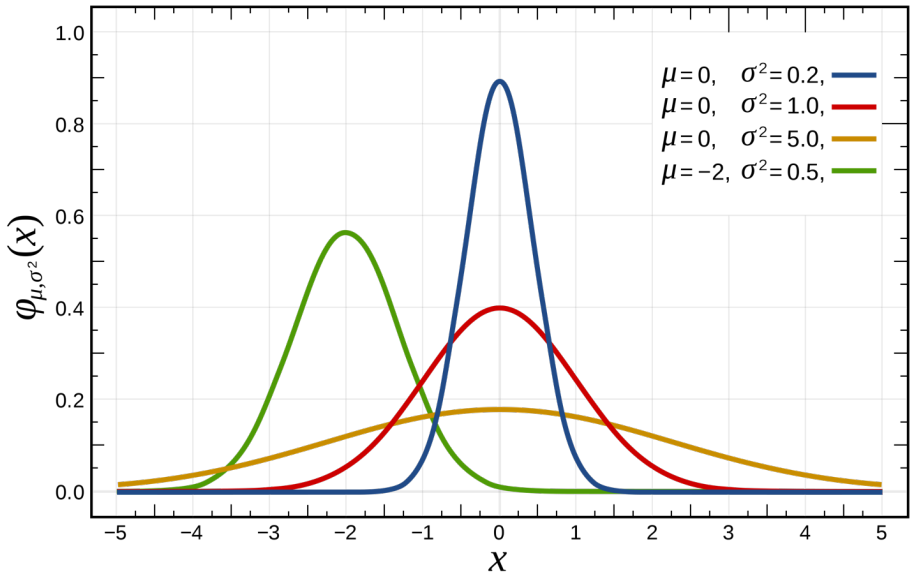


- The *'special'* Autoencoder tries to learn the location of the *'clusters'* instead of the vectors.

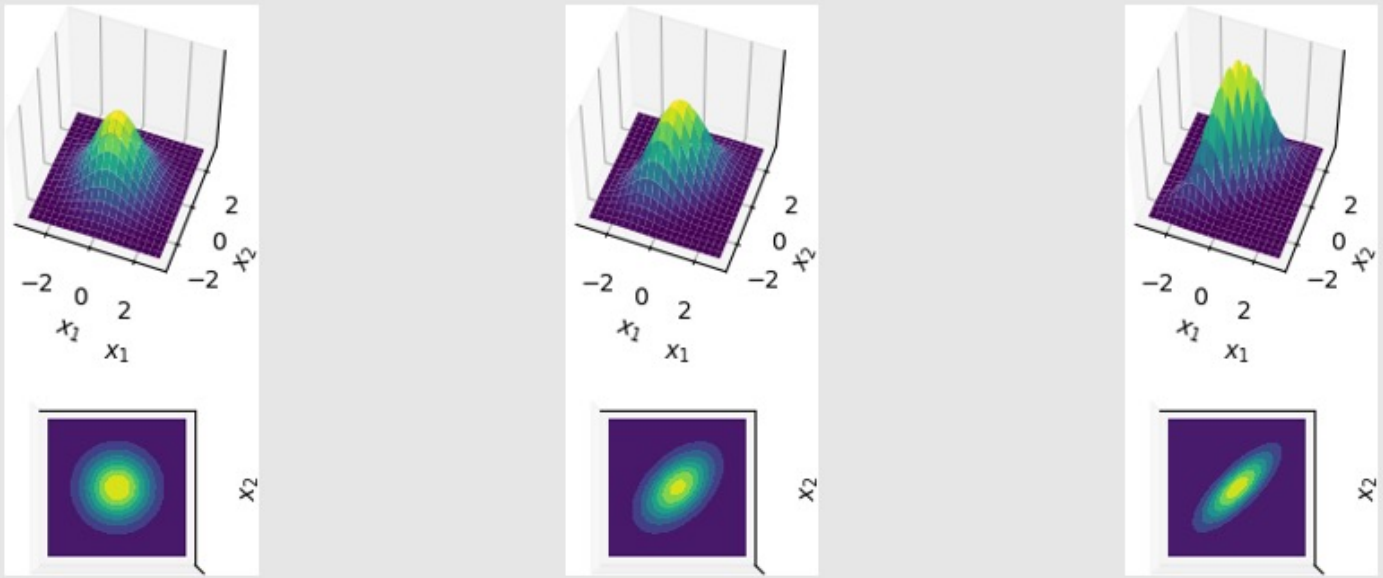
- Learning the clusters:
  - Center of the cluster (μ)
  - Spread of the cluster ($\sum$)

- The *'special'* Autoencoder tries to learn the location of the *'clusters'* instead of the vectors.

- Learning the clusters:
  - Center of the cluster (μ)
  - Spread of the cluster ($\sum$)



$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \qquad \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$



1. Wikipedia, https://en.wikipedia.org/wiki/Gaussian_function

Reference and study material:

1. Deep Learning - Part 1, Dr. Mitesh M. Khapra, NPTEL, https://archive.nptel.ac.in/noc/courses/noc19/SEM2/noc19-cs85/

2. Images are taken from Pixabay, https://pixabay.com/