# Autoencoders
# & Anomaly-detection

Rodrigo Díaz
Manuel Szewc
Ezequiel Alvarez*
sequi@unsam.edu.ar

# Outline

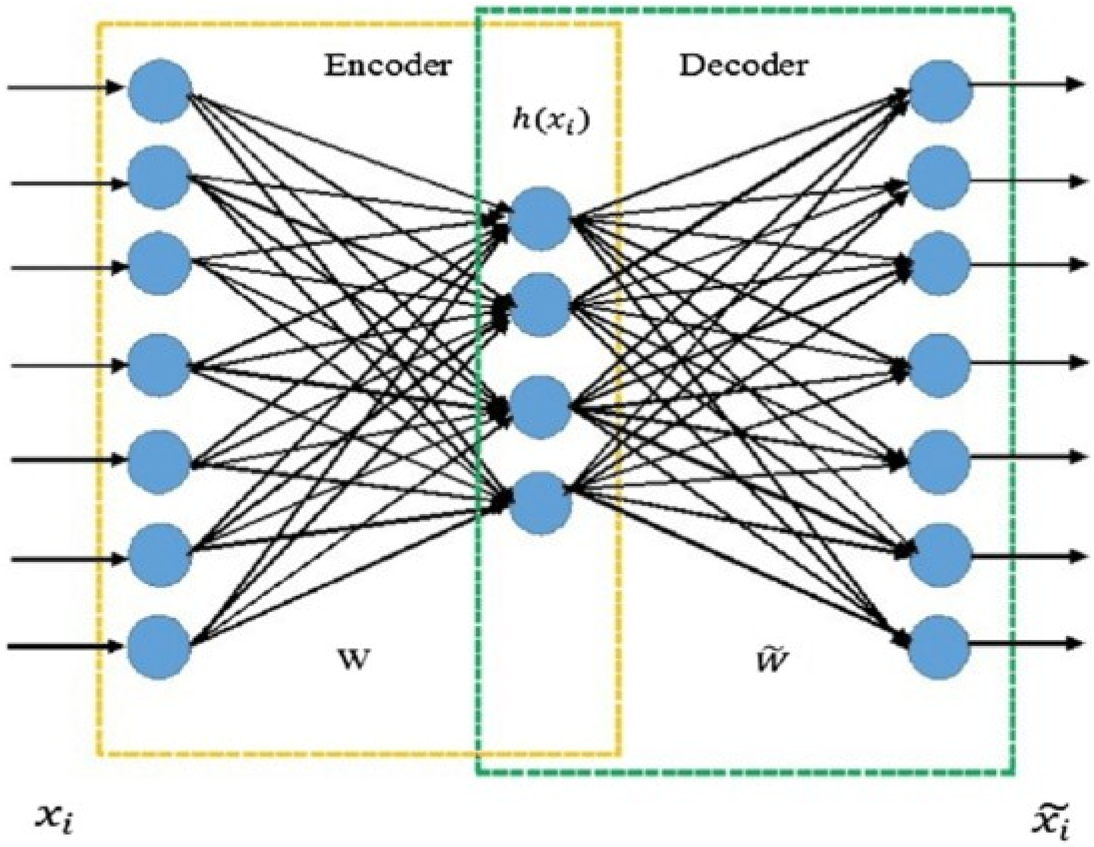
- What is an Auto-Encoder
- Anomaly detection theory
- Notebook 1: Very basic but instructive anomaly example
- Notebook 2: ROC curves on previous example
- Notebook 3: More sophisticated example
- Notebook 4: Credit Card fraud detection
- Intro to Variational Auto-Encoders (VAE)

**Hands-on!**

Credits / bibliography:


https://towardsdatascience.com/a-keras-based-autoencoder-for-anomaly-detection-in-sequences-75337eaed0e5
https://www.jeremyjordan.me/variational-autoencoders/
https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf
https://medium.com/@realityenginesai/understanding-variational-autoencoders-and-their-applications-81a4f99efc0d

# Auto-Encoder

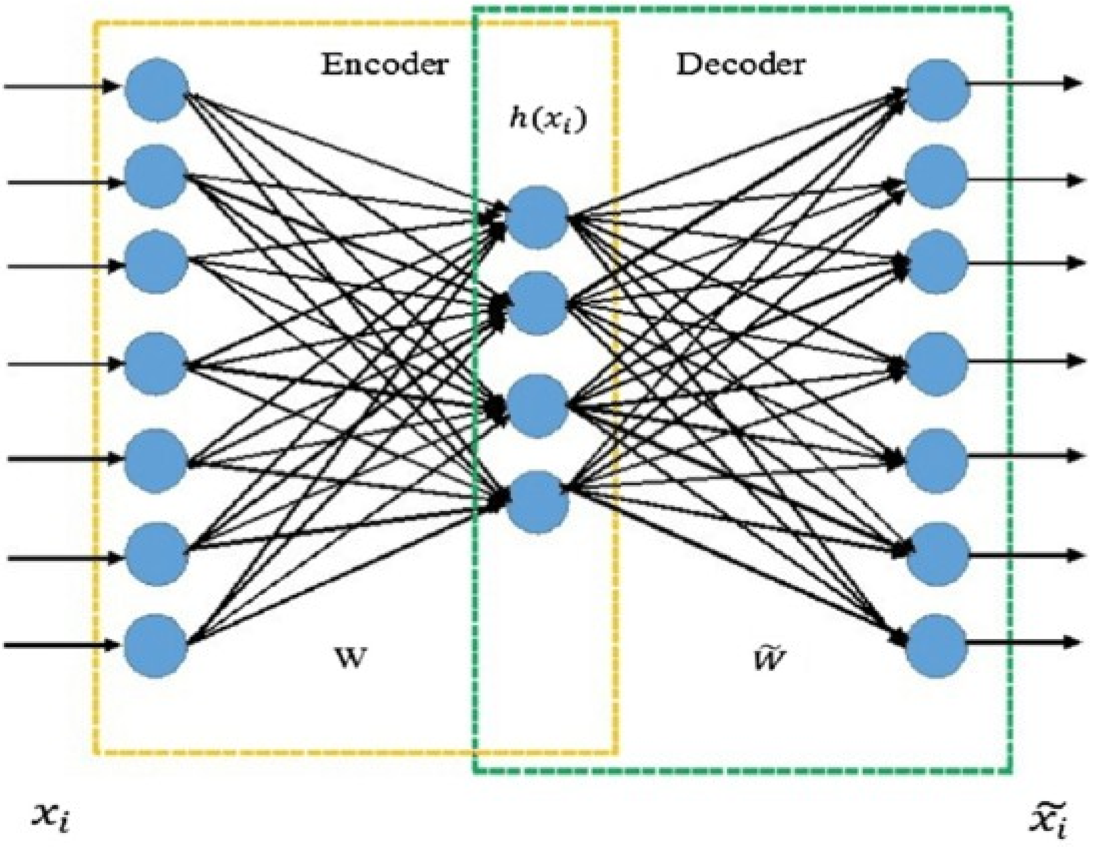



- It is a normal NN
- Specific topological architecture
- Unsupervised:

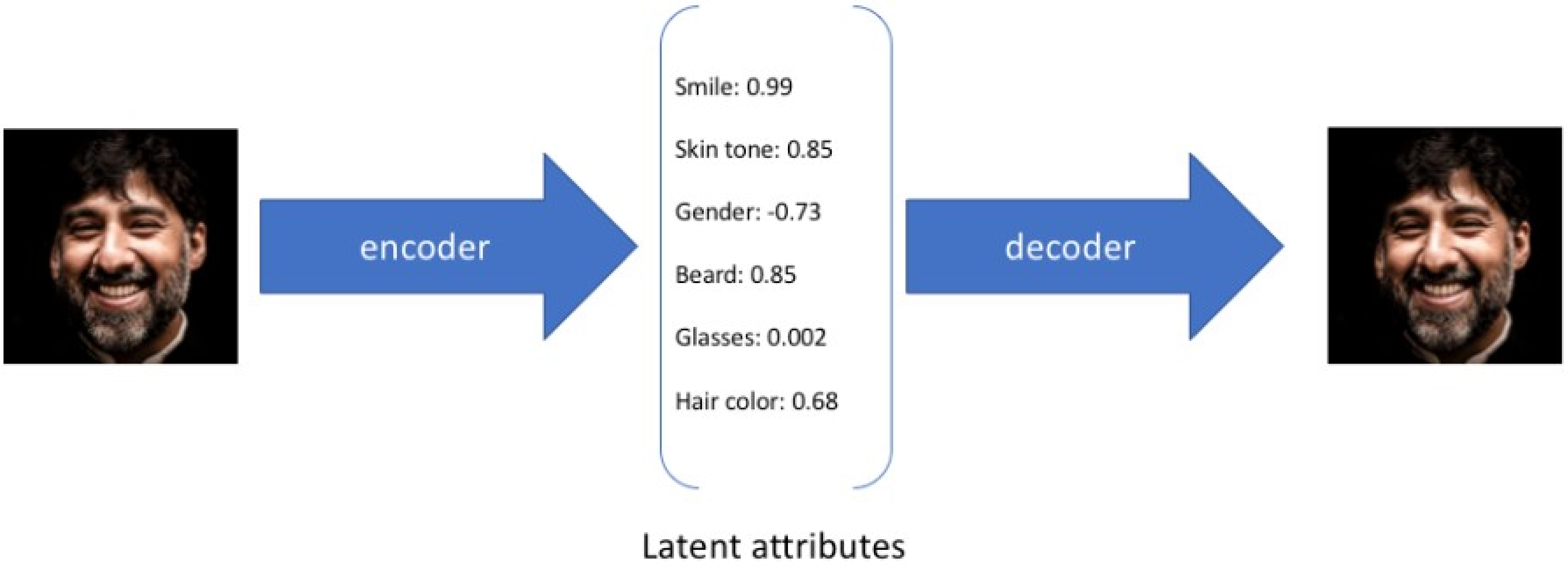
  request OUT=IN

# Auto-Encoder

- Encoder → Latent → Decoder
- Latent: Reduce info
- NN: minimizing loss-function is equivalent to **selecting relevant features**



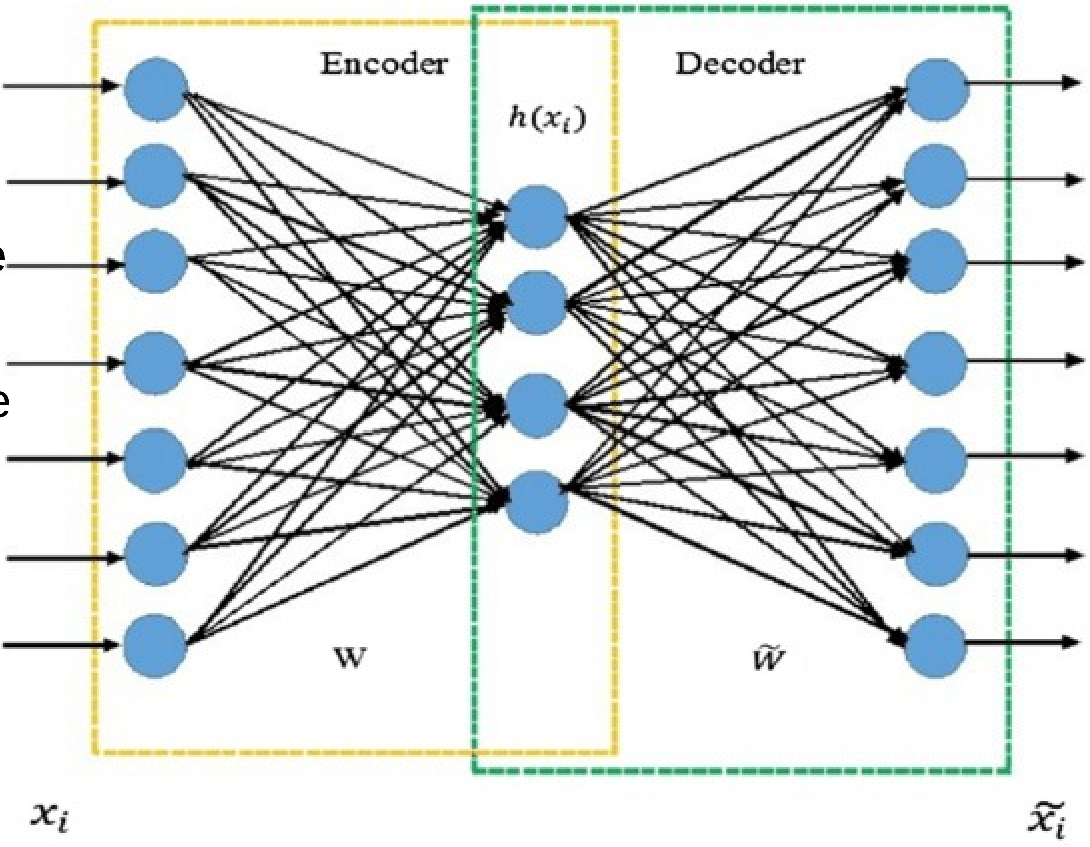
# Auto-Encoder



# Auto-Encoder for anomalies

- How to use it to detect anomalies?
- Latent space: get the relevant features of the **majority** of the events
- When attempting to reproduce an anomaly will fail
- Mean Square Error (MSE)

  sqrt( ( x1-y1 )² + ( x2-y2 )² +..)

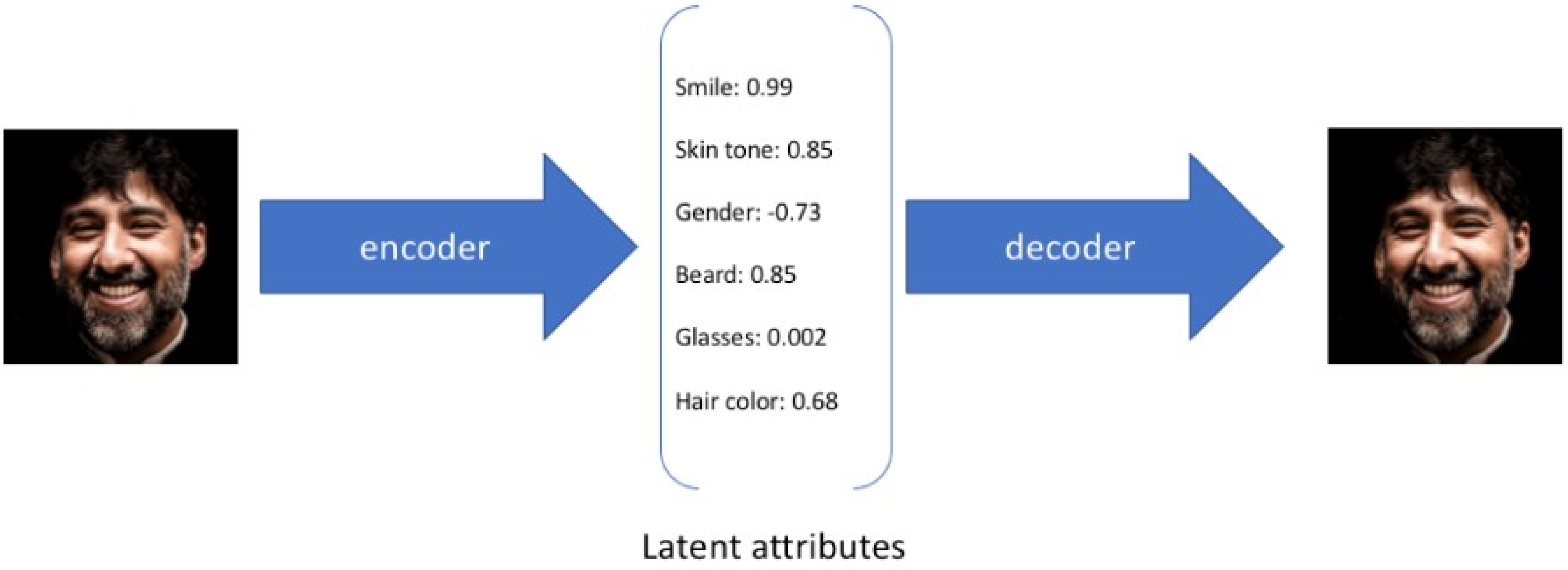


# Let’s go with the notebooks

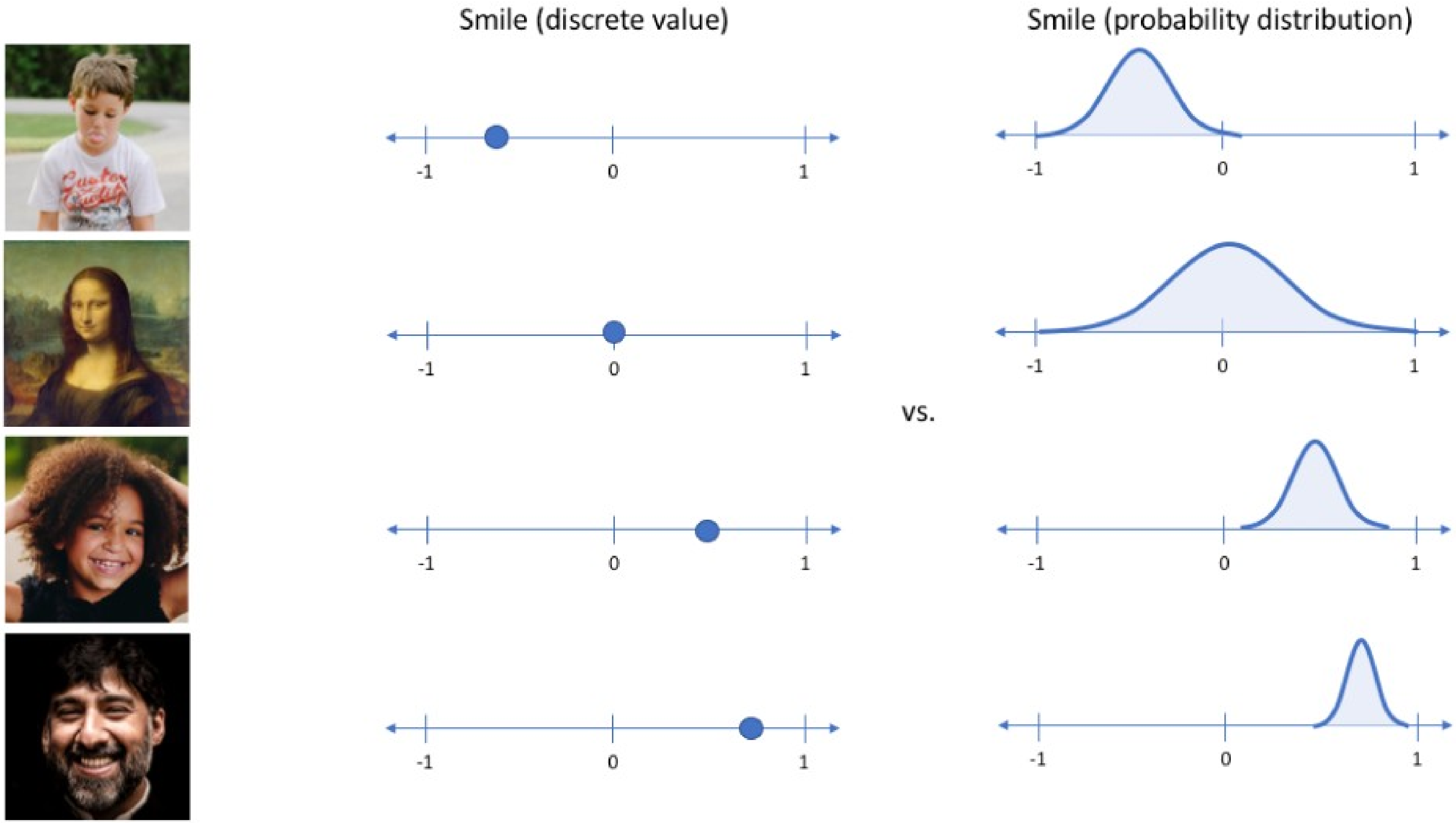
**Hands-on!**

Notebook 1: Very basic but instructive anomaly example
Notebook 2: ROC curves on previous example
Notebook 3: More sophisticated example
Notebook 4: Credit Card fraud detection

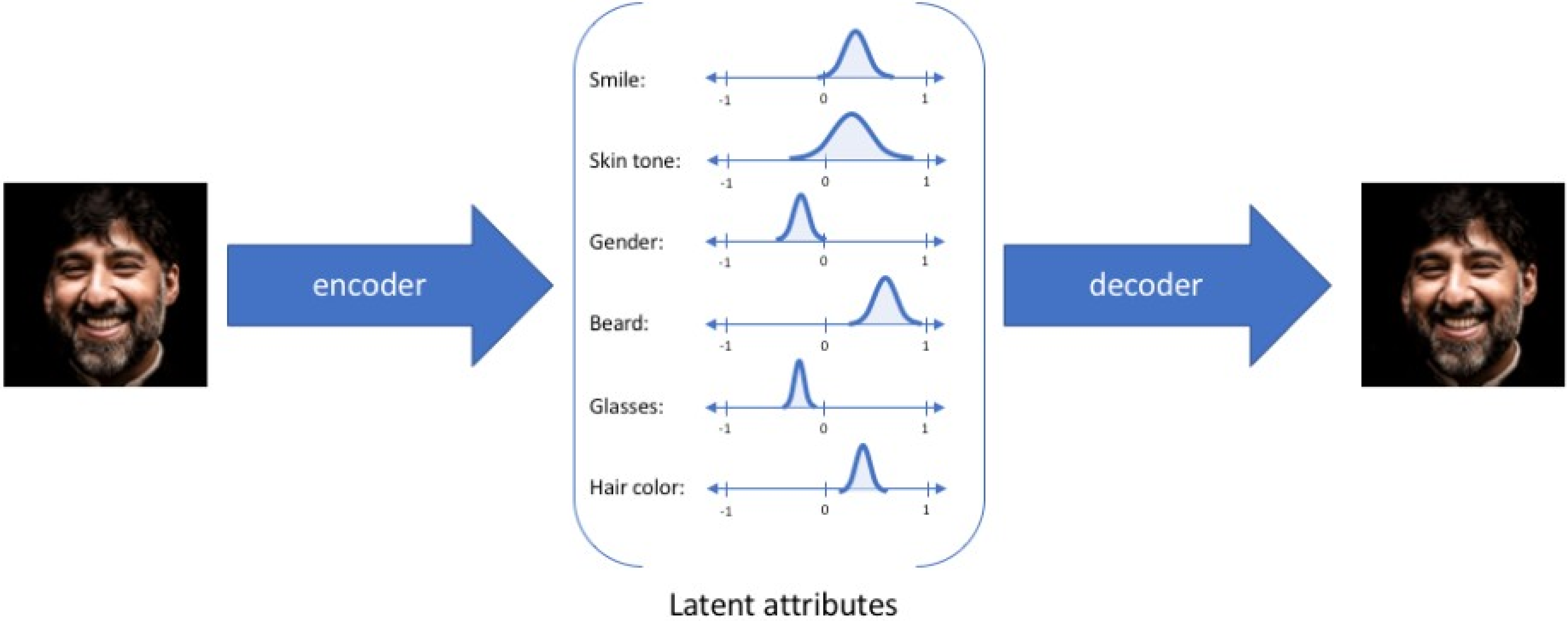
# Variational Auto-Encoders (intro)



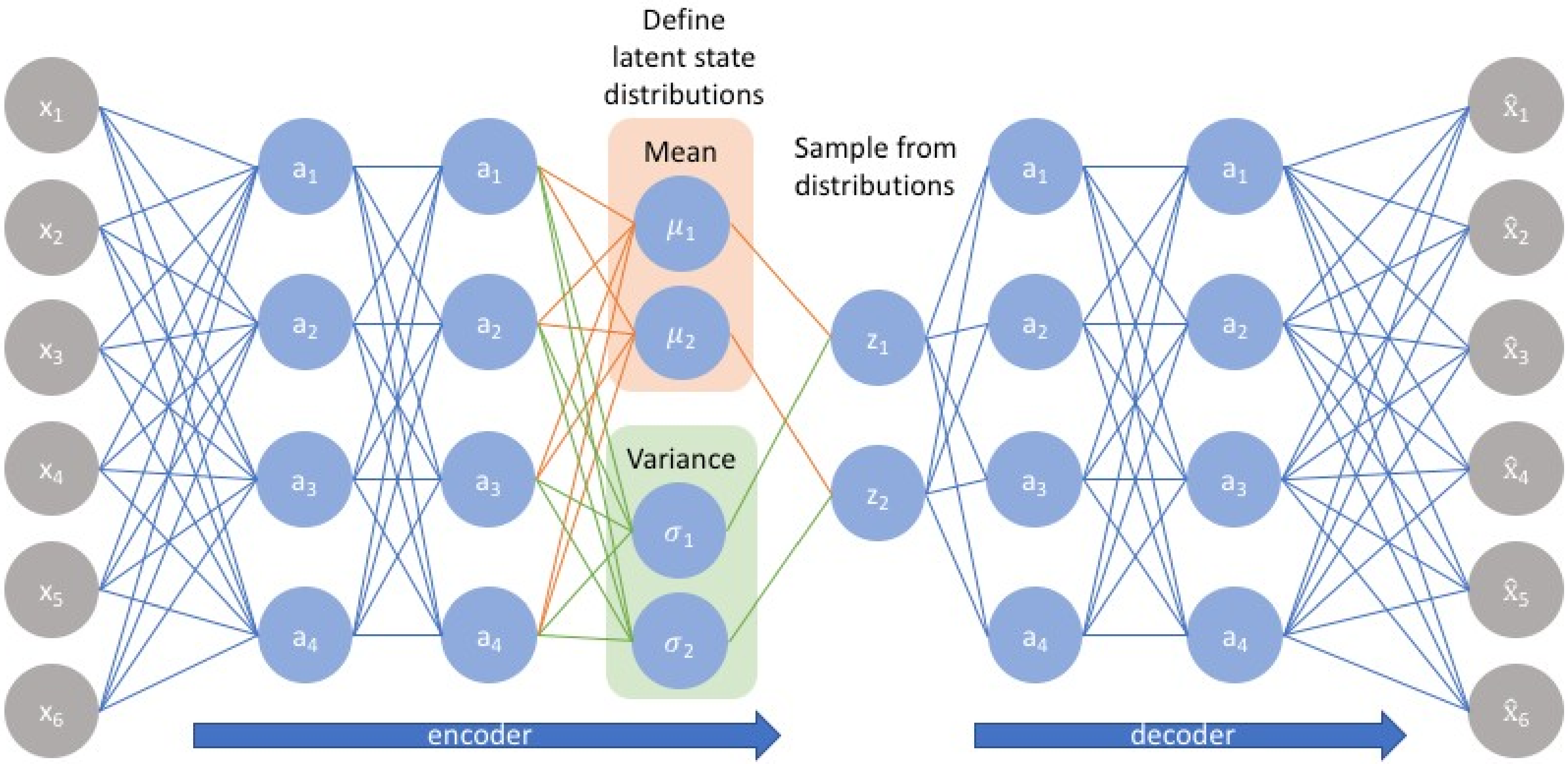
# Variational Auto-Encoders (intro)



# Variational Auto-Encoders (intro)

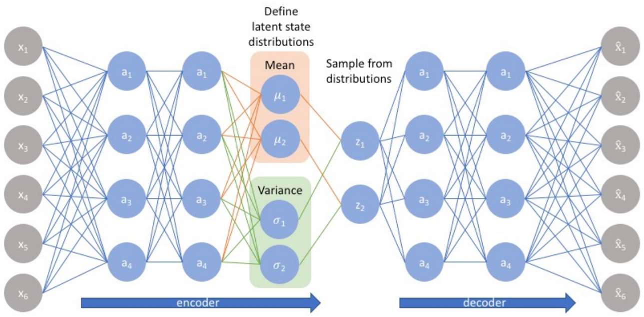


# Variational Auto-Encoders (intro)



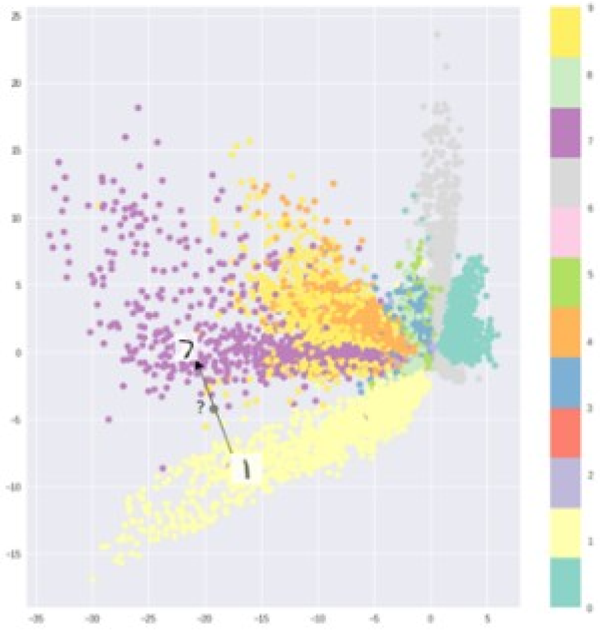
# Variational Auto-Encoders (intro)

- AEs cluster data-points in Latent Space → any random sampling in Latent Space may produce not existing outputs
- VAEs fix this issue since the latent space is now sampled from the PDF specified by the “pre-Latent” Space.
- Price to pay: you need more data to correctly train a VAE
- IMO: AEs are better when data is not huge, VAEs are better with huge data, and VAEs are crucial for sampling senseful outputs through the Auto-Encoder



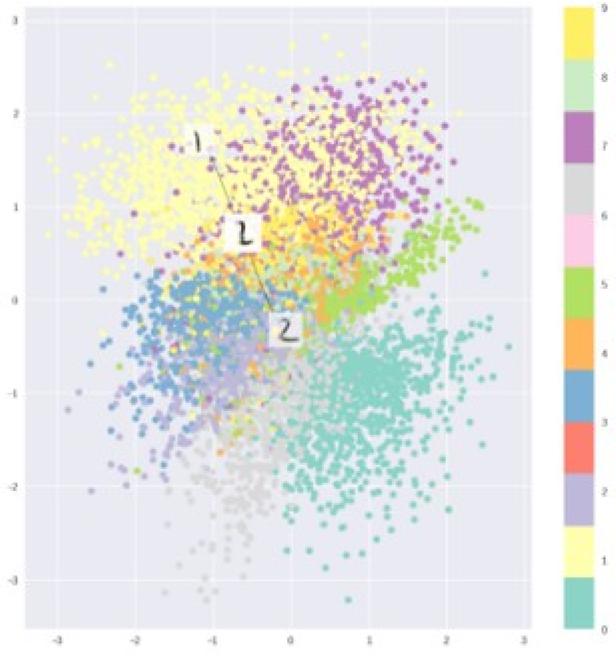
# Variational Auto-Encoders (intro)

MINST dataset with 2D Latent Space:

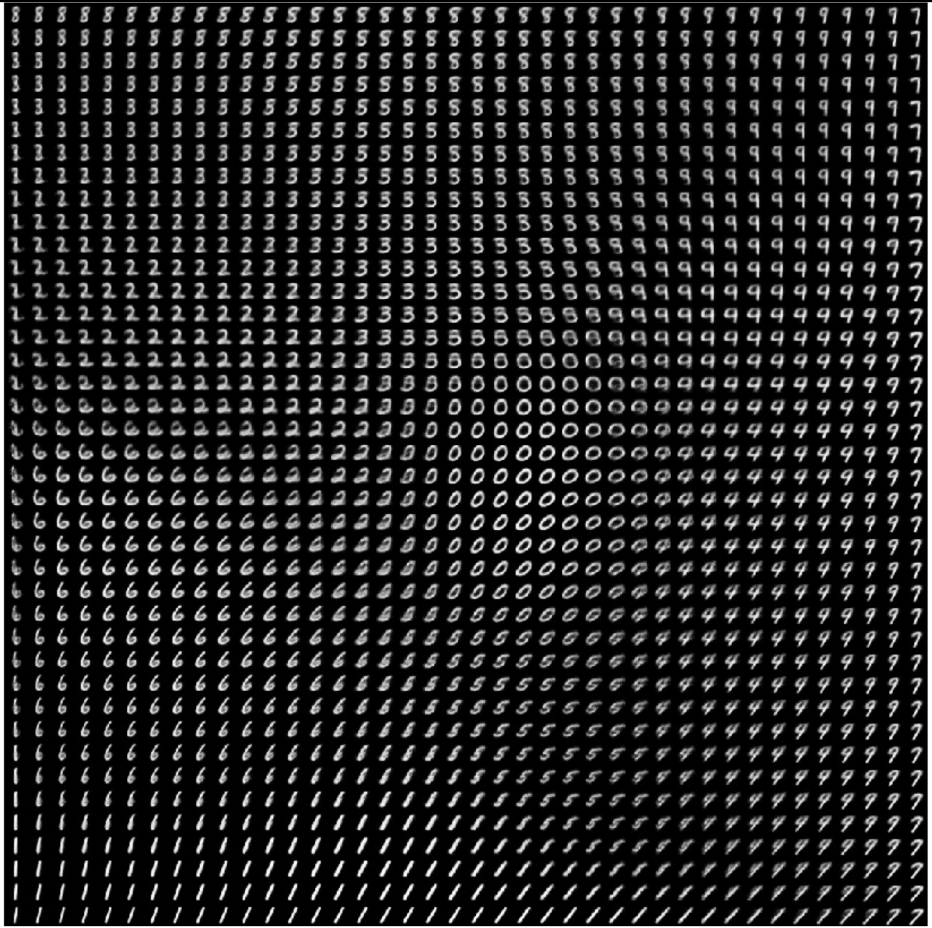


← Auto-Encoder

Variational
Auto-Encoder →



# Variational Auto-Encoders (intro)



Data generated by the decoder network of a VAE trained on the MINST dataset.

Here, it is sampled a grid of values from a two-dimensional Gaussian and displayed the output of the decoder network.

# Summary & Conclusions

**Auto-encoders:**

- NN with a very specific shape
- Doesn’t need labeling
- Unsupervised

**Notebooks:**

- Hands-on academic example
- We worked an ‘anomaly-study’ though fictitious anomalies sampling
- Get some insight on architecture
- Real data: Credit Card fraud

**Variational Auto-encoders:**

- Replace LS → sampling
- Smoother LS
- Generative model

# Summary & Conclusions

**Auto-encoders:**

- NN with a very specific shape
- Doesn't need labeling
- Unsupervised

**Notebook**

## Muchas Gracias!

ugh

- Real data: Credit Card fraud

**Variational Auto-encoders:**

- Replace LS → sampling
- Smoother LS
- Generative model