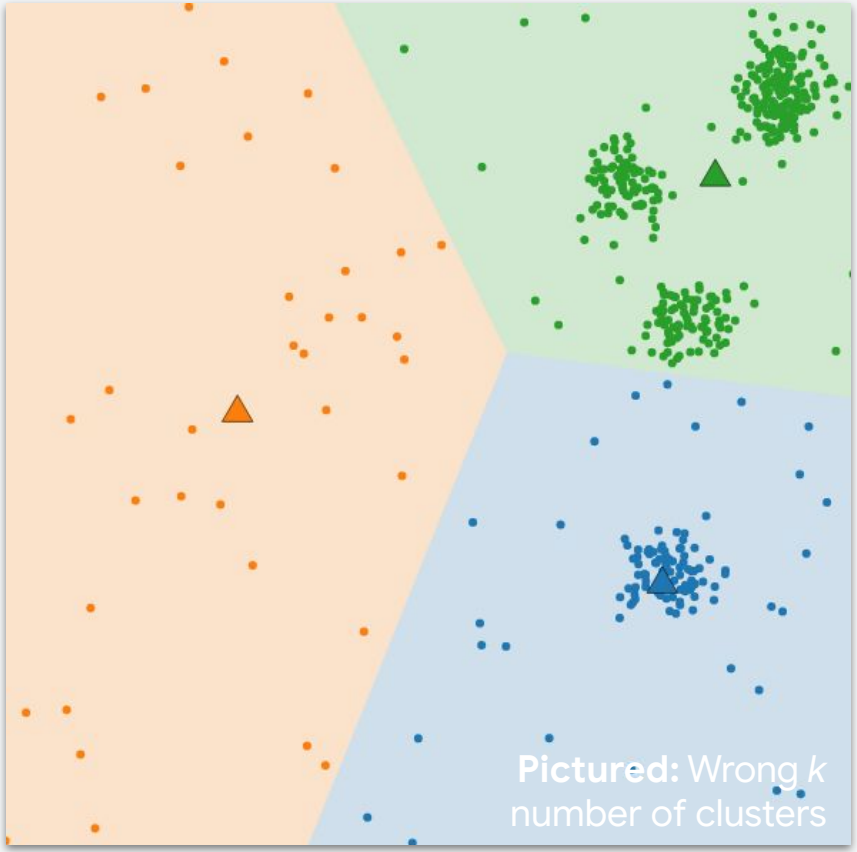
Autoencoders

Recap of **K-Means**

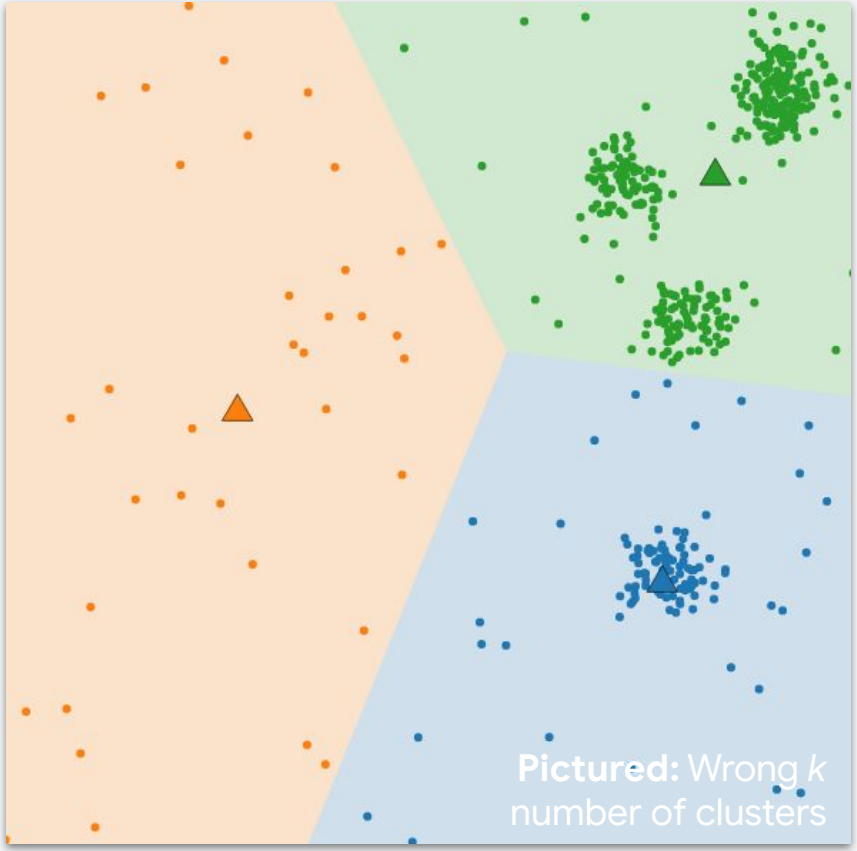
Goal:

Partition *n* observations into *k* clusters



When **K-Means** Fails?

- Scales poorly with **large numbers** of observations
- **“Curse of Dimensionality”**

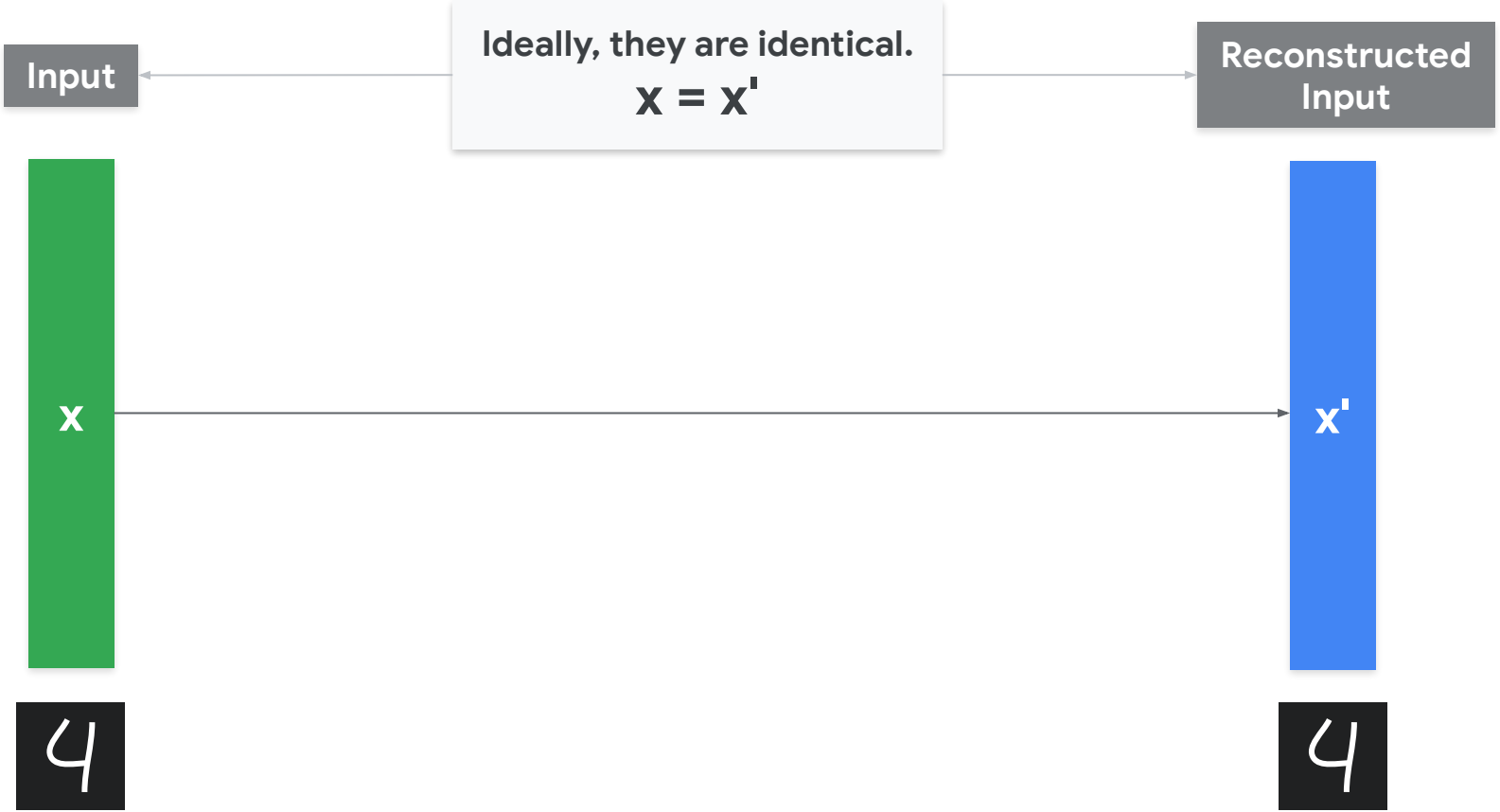


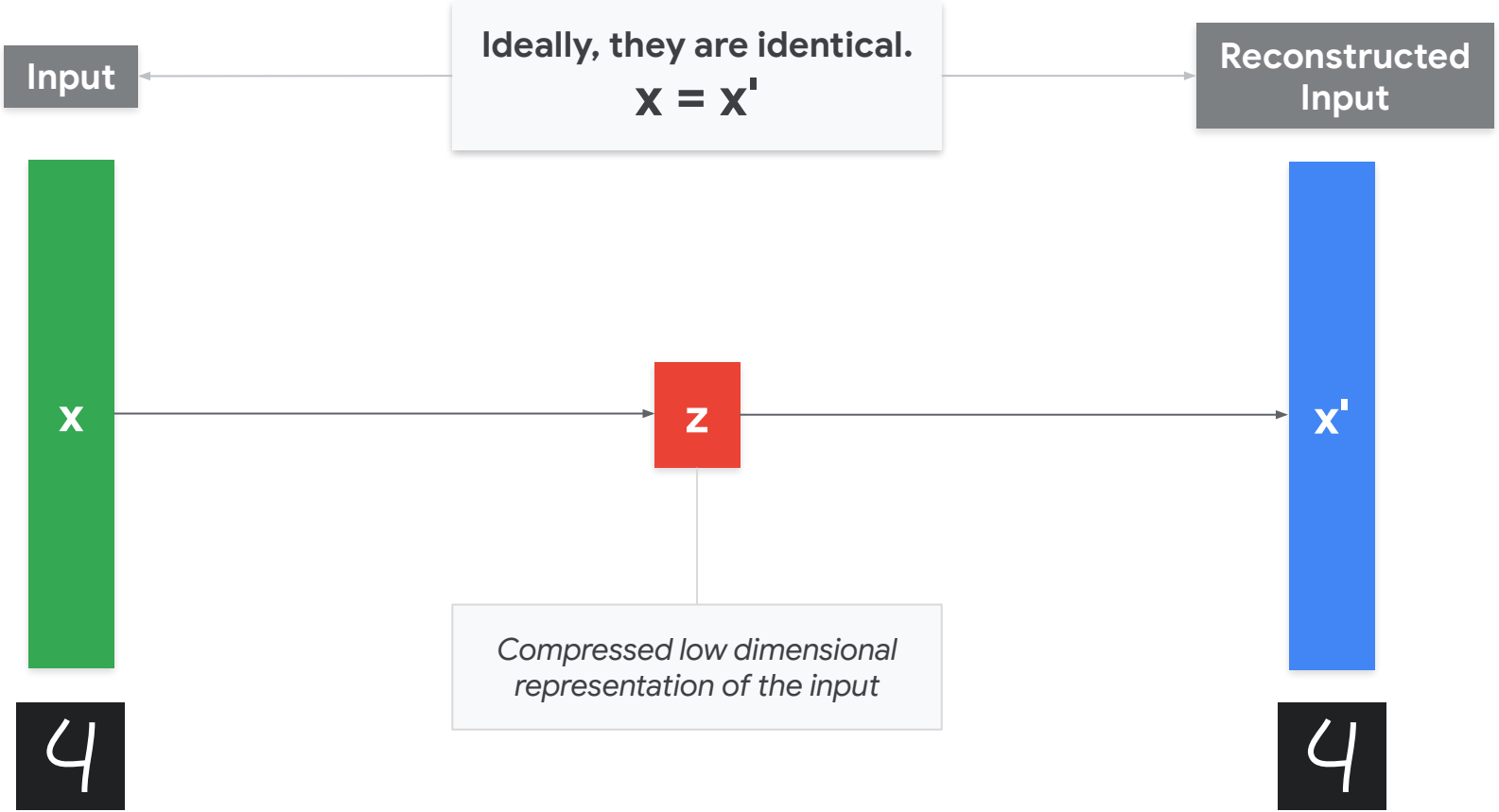

Input
Prediction
Error

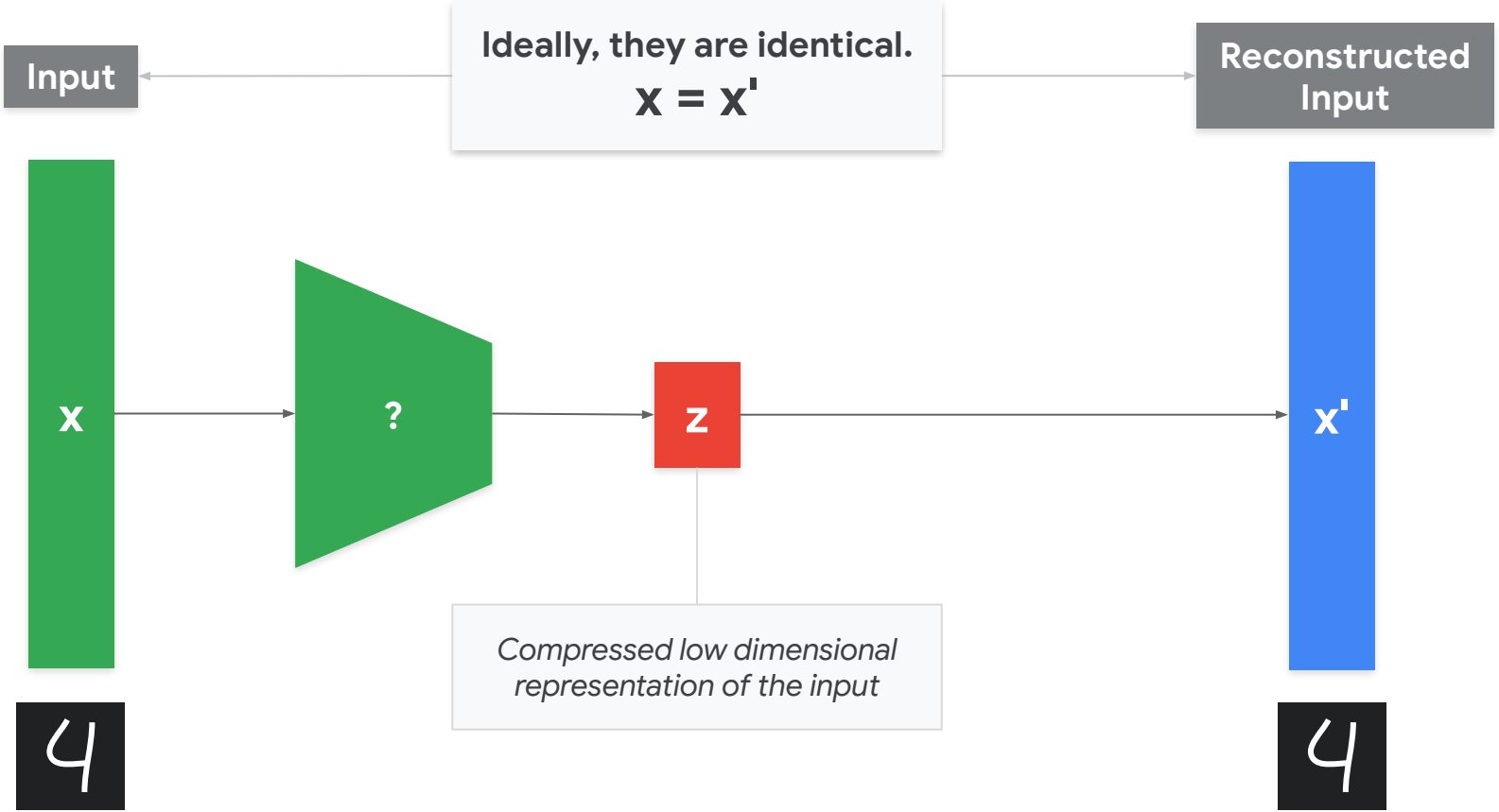
You don’t know what you **don’t know**.

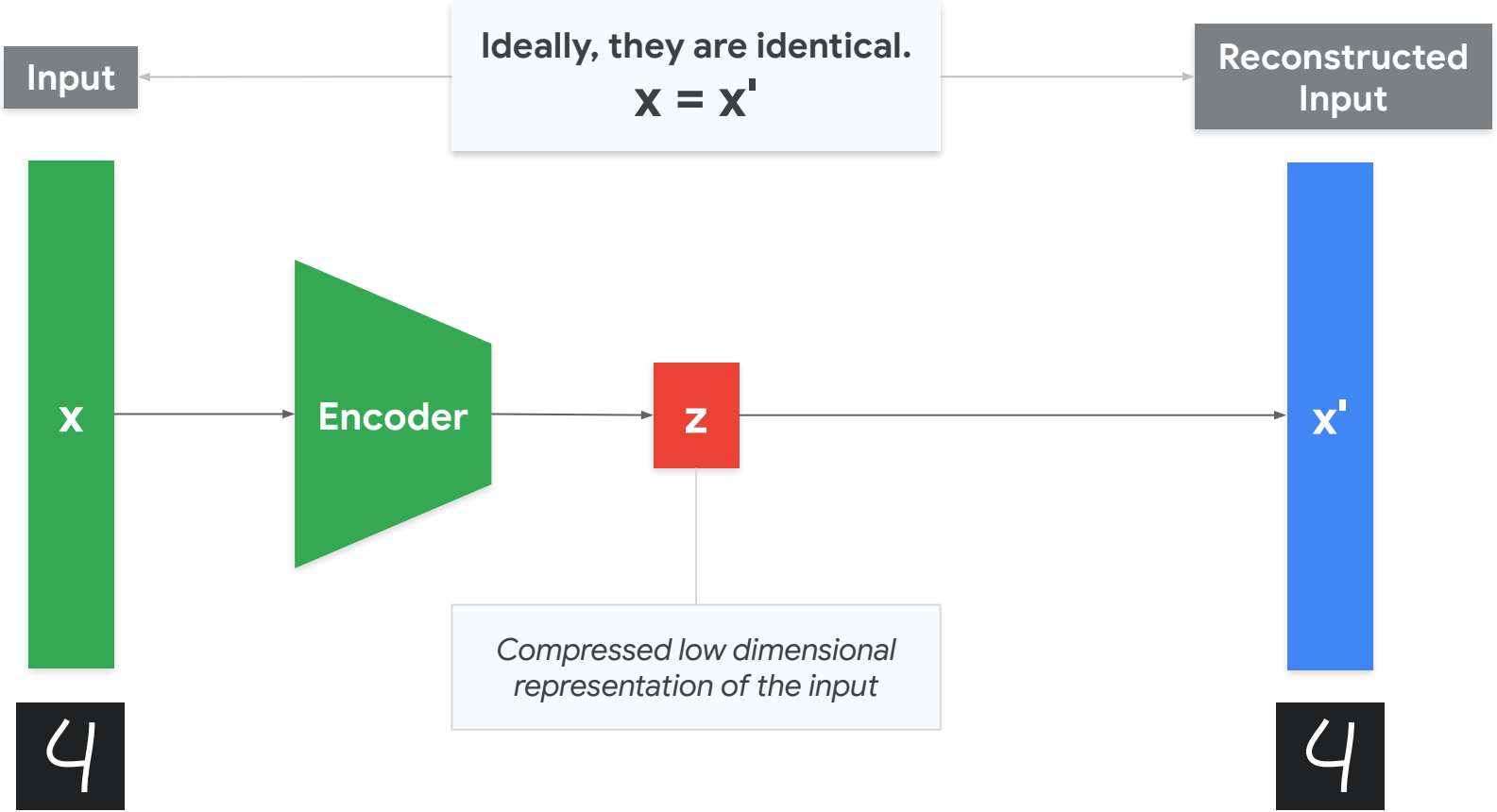
You don't know what you **don't know**.

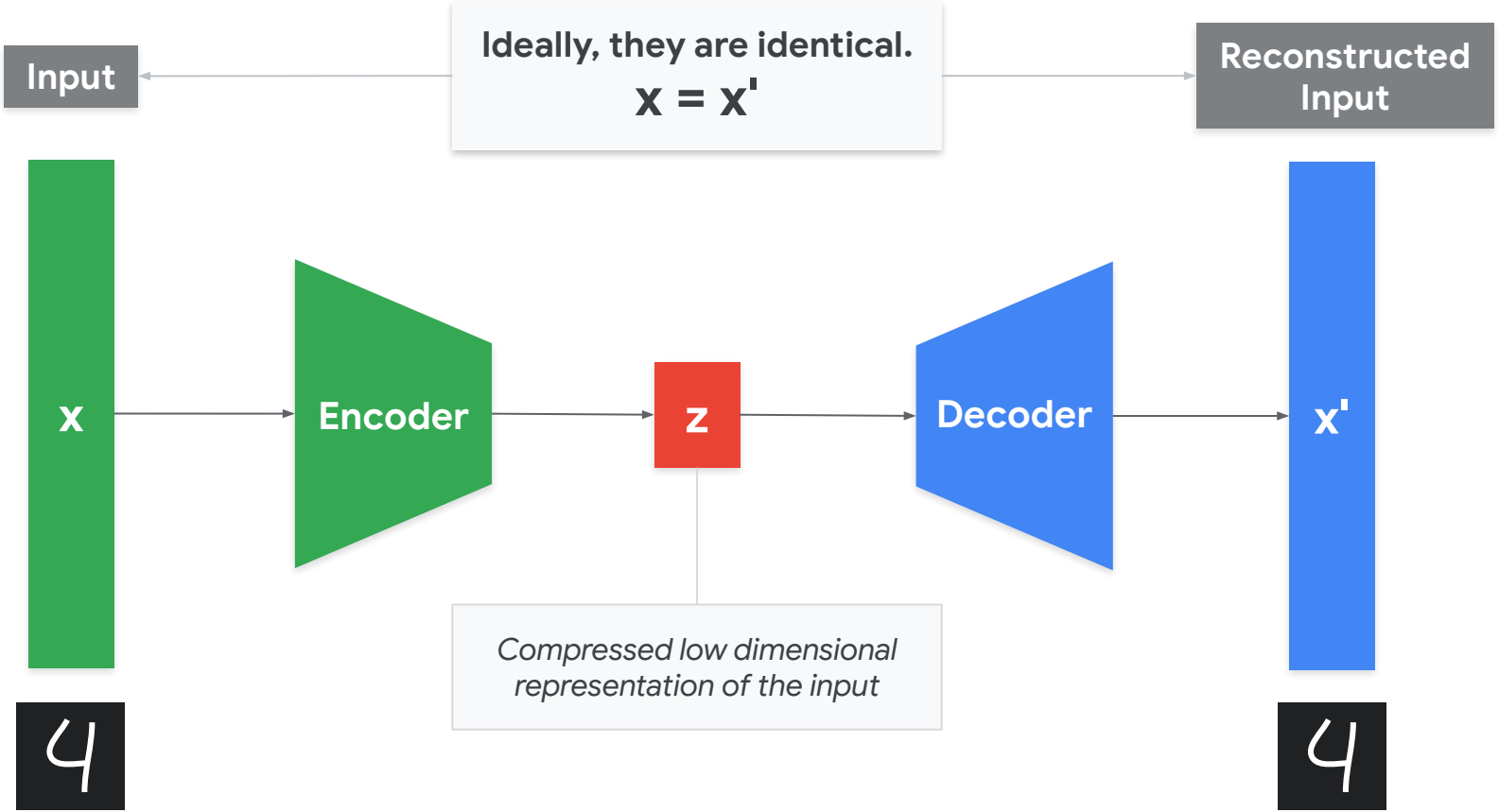
But **you do know** what **you know**.

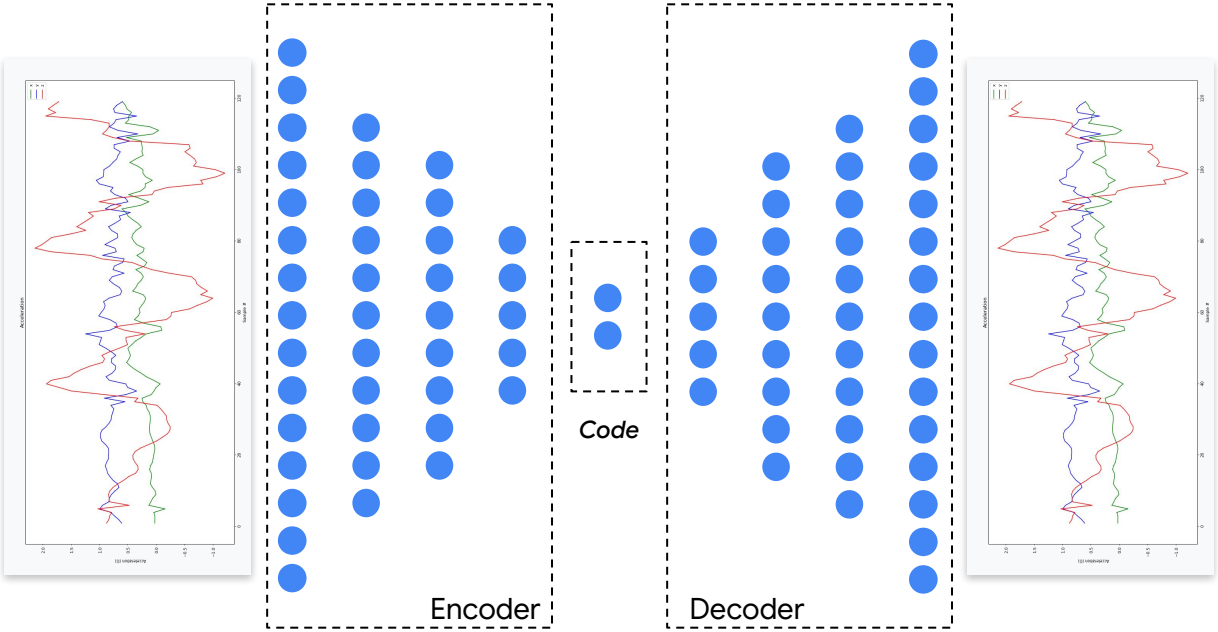

Input
Ideally, they are identical.
x = x'
Reconstructed Input
x
x'

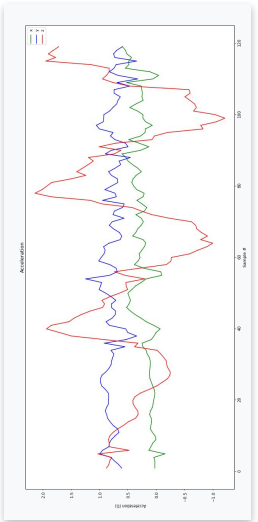

Input
Ideally, they are identical.
x = x'
Reconstructed Input
x
z
x'
Compressed low dimensional representation of the input

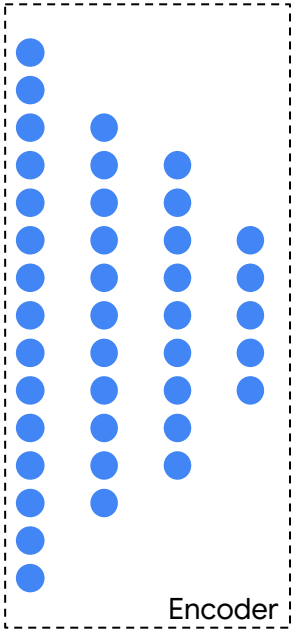

Input
Ideally, they are identical.
x = x'
Reconstructed Input
x
?
z
x'
Compressed low dimensional representation of the input

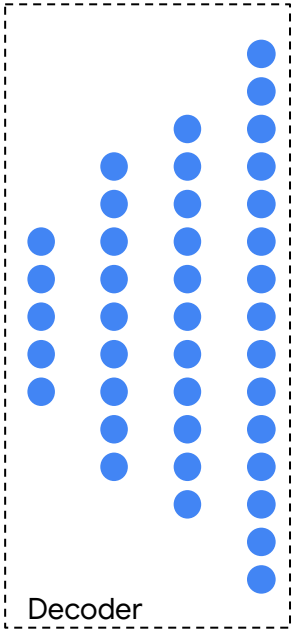

Input
Ideally, they are identical.
x = x'
Reconstructed Input
x
Encoder
z
x'
Compressed low dimensional representation of the input

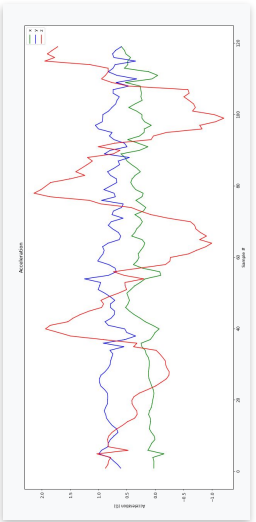

Input
Ideally, they are identical.
x = x'
Reconstructed Input
x
Encoder
z
Decoder
x'
Compressed low dimensional representation of the input


Encoder
Code
Decoder

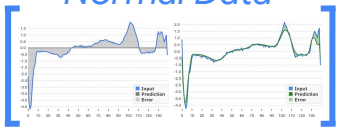




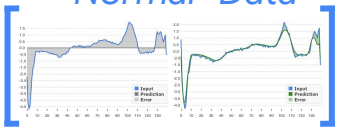


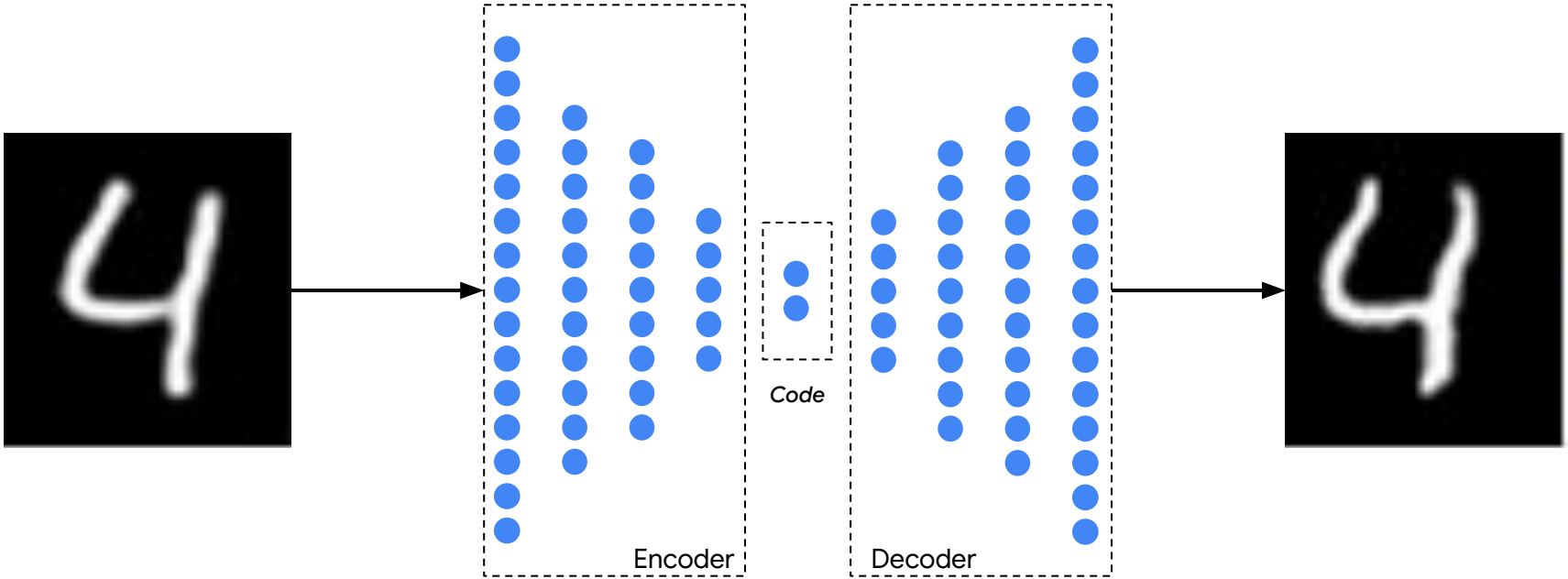


Normal Data

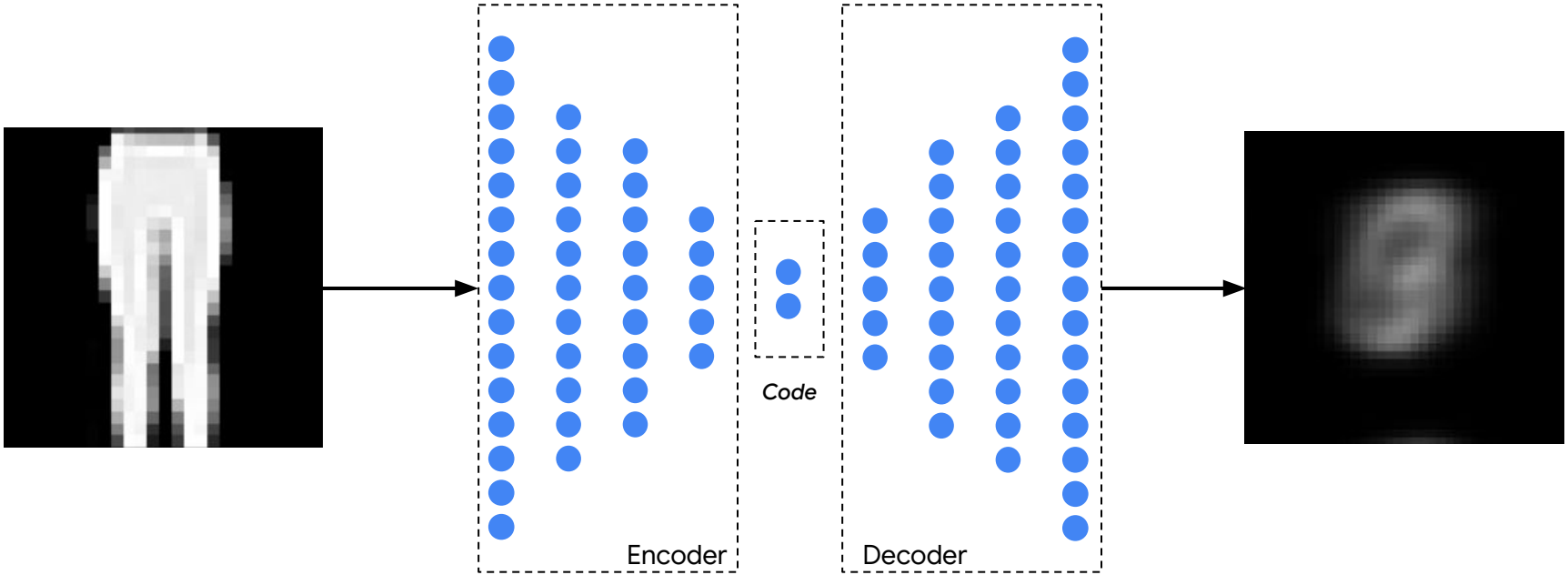


"Normal" Data



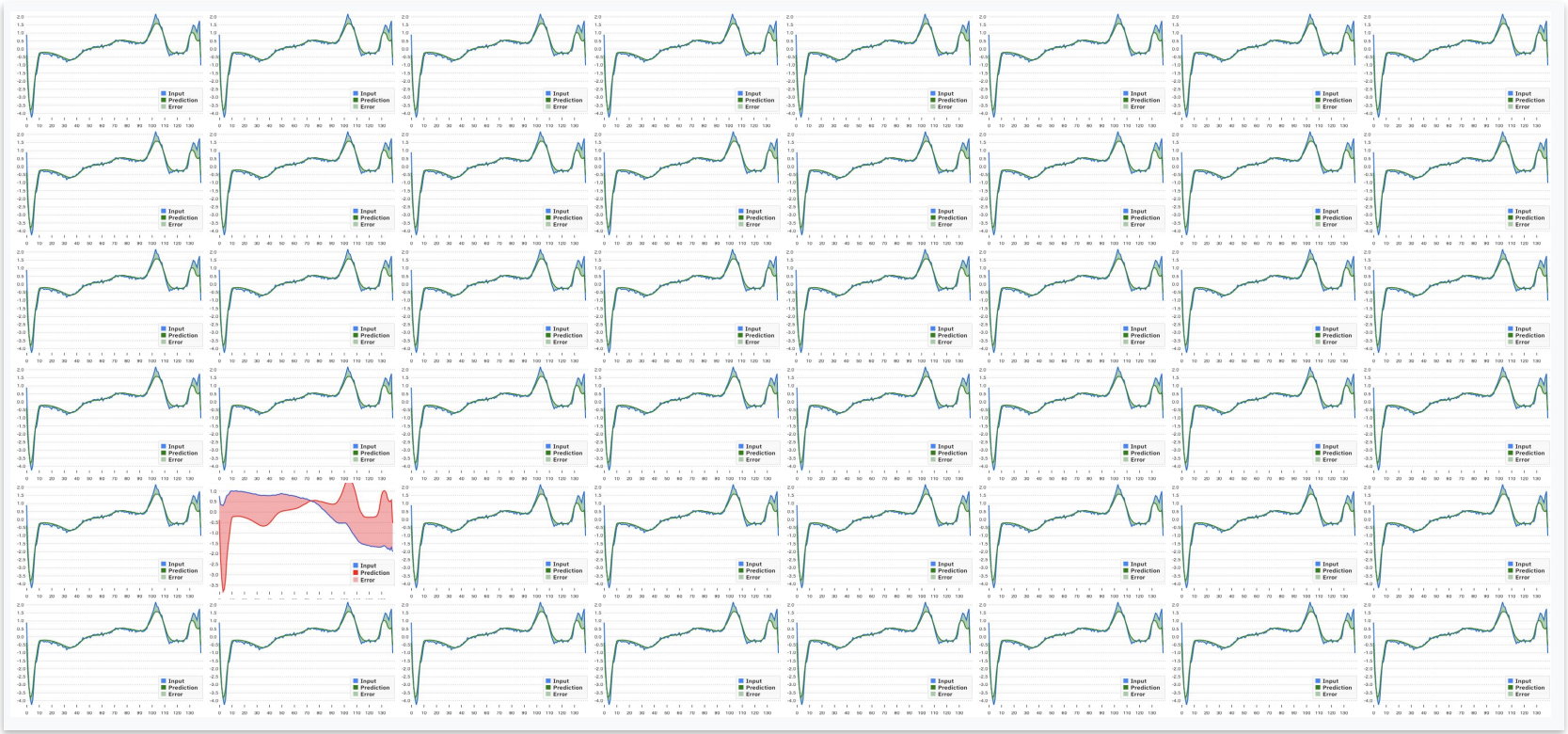


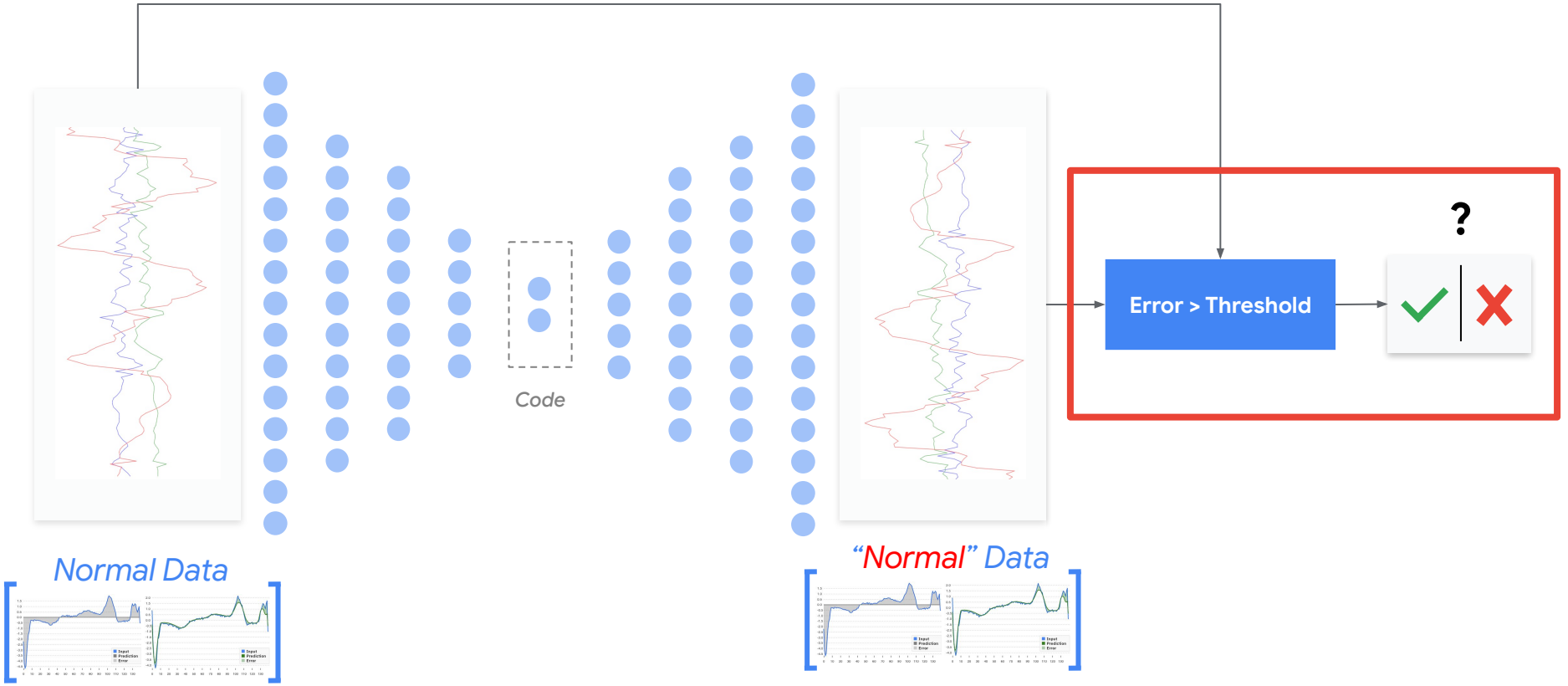
Low Reconstruction Loss
(Low Mean Square Error)



High Reconstruction Loss
(High Mean Square Error)

Unbalanced Data




Code
Error > Threshold
?
Normal Data
"Normal" Data