[A4-004] 딥러닝 코딩 실습

Lecture 01: Auto-Encoder

Hak Gu Kim

hakgukim@cau.ac.kr

Immersive Reality & Intelligent Systems Lab (IRIS LAB)

Graduate School of Advanced Imaging Science, Multimedia & Film (GSAIM)

Chung-Ang University (CAU)

05 Jan. 2023

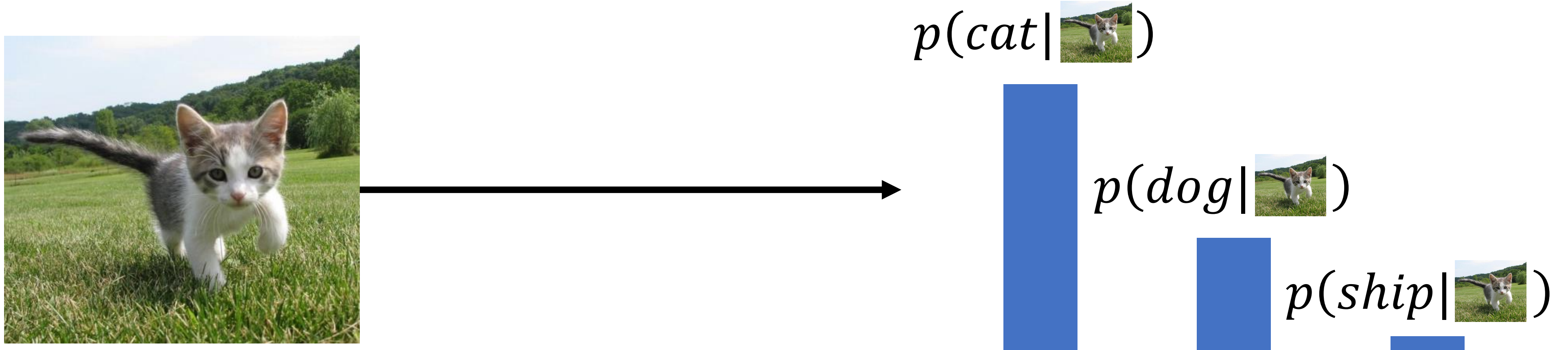
Topic

- Generative Models
- Auto-Encoders

***Note:** Many of these slides in this course were adapted from Convolutional Neural Networks for Visual Recognition (Stanford Univ.) and Deep Learning for Computer Vision (Univ. of Michigan)

Discriminative Model vs. Generative Model

- Discriminative Model
  - Learn a probability distribution $p(y|x)$



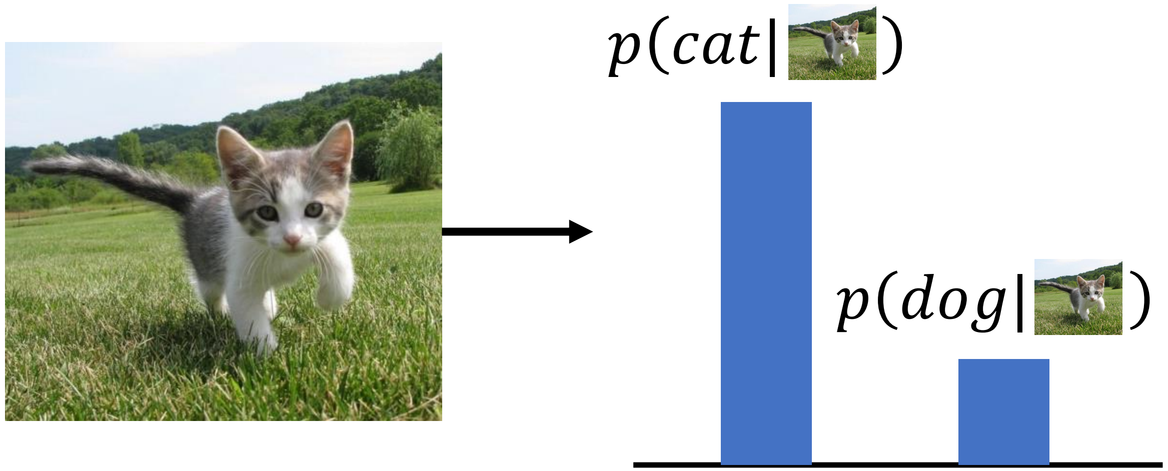
Data x

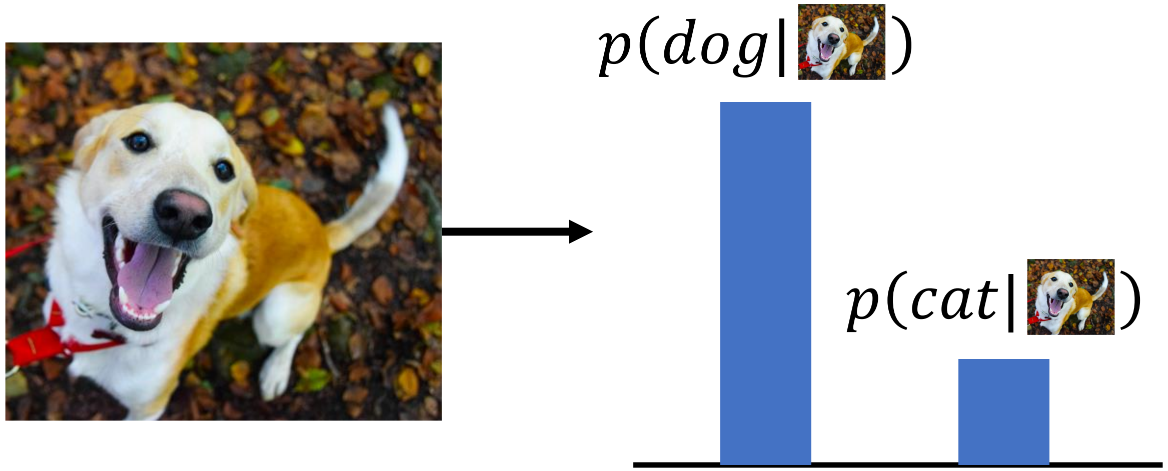
Discriminative Model vs. Generative Model

- Discriminative Model

— Learn a probability distribution $p(y|x)$

— The possible labels for each input compete for probability mass. But no competition between images





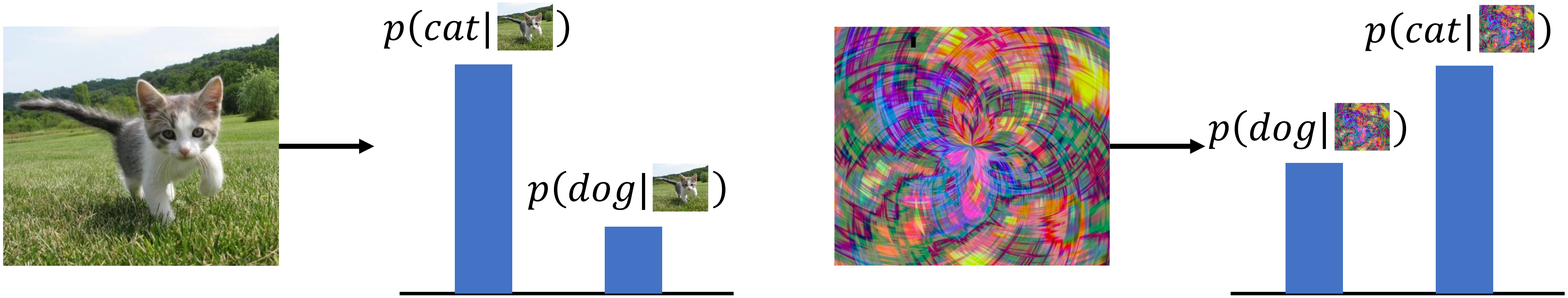
Discriminative Model vs. Generative Model

- Discriminative Model

— Learn a probability distribution $p(y|x)$

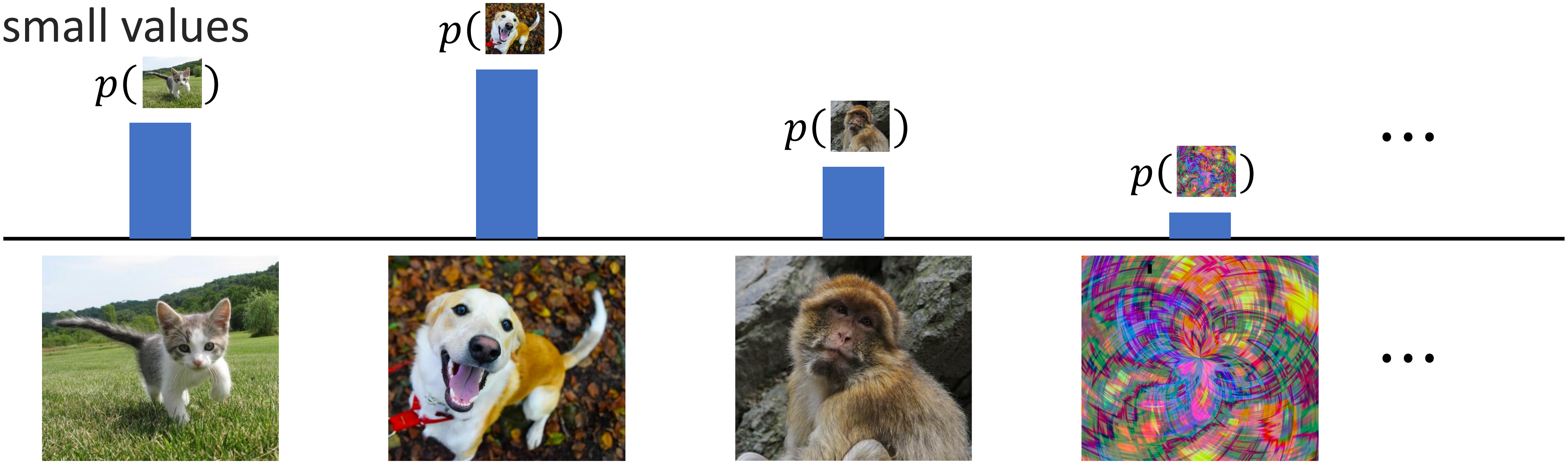
— The possible labels for each input compete for probability mass. But no competition between images

— No way for the model to handle unreasonable inputs. It must give label distributions for all images



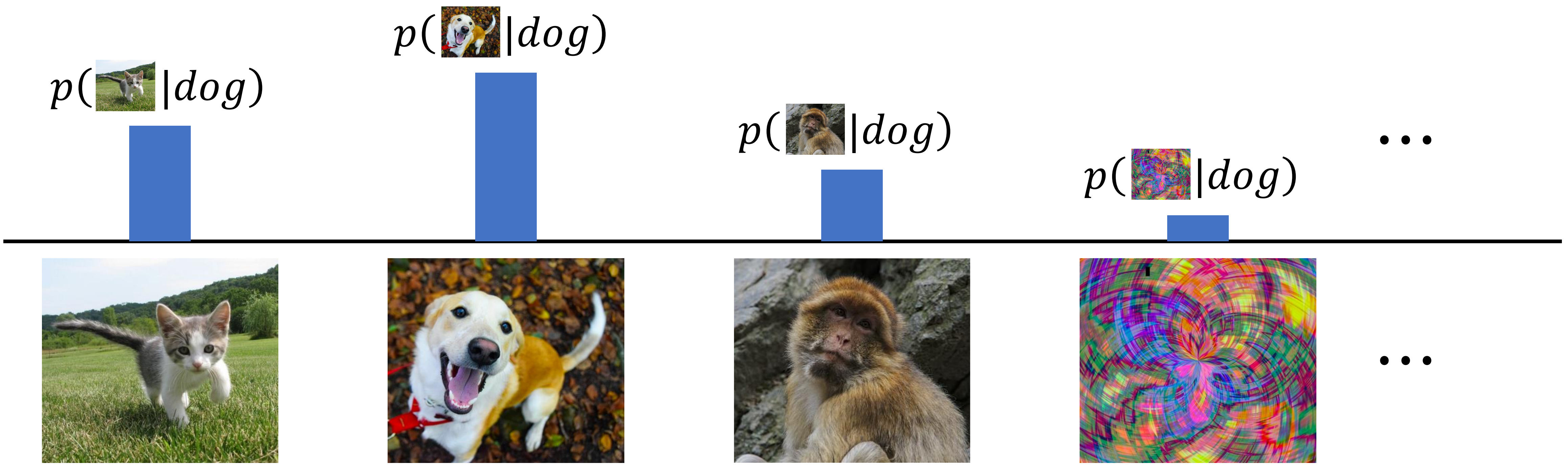
Discriminative Model vs. **Generative Model**

- Generative Model
  — Learn a probability distribution $p(x)$
  — All possible images compete with each other for probability mass
  — The generative model can reject unreasonable inputs by assigning them small values



Discriminative Model vs. **Generative Model**

- **Conditional** Generative Model
  — Learn a probability distribution $p(x|y)$
  — Each possible label induces a competition among all images



Discriminative Model vs. Generative Model

- Discriminative Model

— Learn a probability distribution $p(y|x)$

- Generative Model

— Learn a probability distribution $p(x)$

- Conditional Generative Model

— Learn a probability distribution $p(x|y)$

Discriminative Model

Conditional Generative Model

$$P(x|y) = \frac{P(y|x)}{P(y)} P(x)$$

Prior over labels

(Unconditional) Generative Model

Applications of Each Model

- Discriminative Model

— Learn a probability distribution $p(y|x)$

- ✓ Assign labels to data
- ✓ Feature learning with labels

- Generative Model

— Learn a probability distribution $p(x)$

- ✓ Outlier detection
- ✓ Feature learning without labels
- ✓ Sample to generate new data

- Conditional Generative Model

— Learn a probability distribution $p(x|y)$

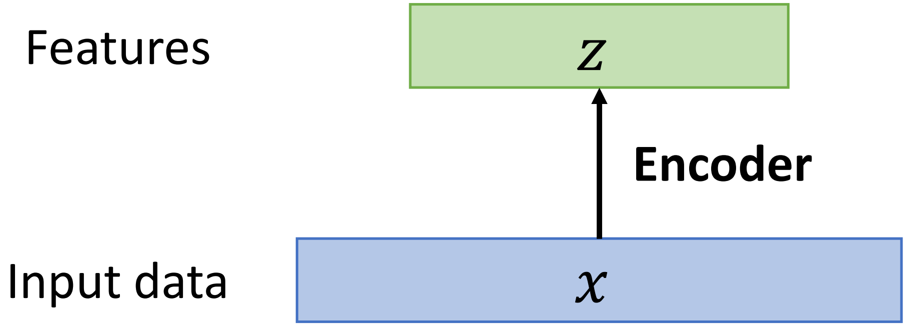
- ✓ Assign labels while rejecting outliers
- ✓ Generate new data conditioned on labels

Auto-Encoders

- An unsupervised method is to learn feature vectors from raw data x without any labels y

- How can we learn the feature transform from raw data?

— Feature should extract useful information (e.g., object identities, scene type, *etc.*) that we can use for downstream tasks

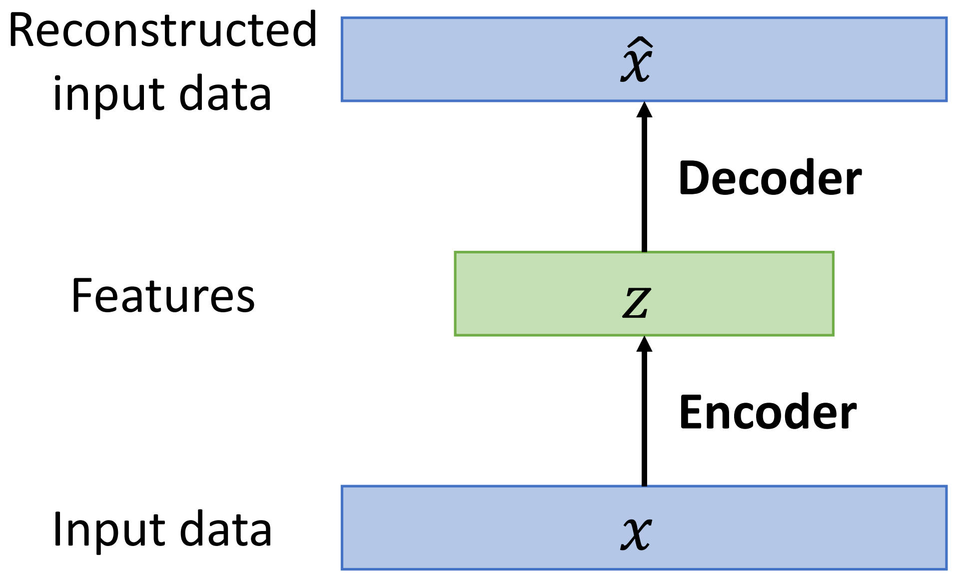


Auto-Encoders

- How can we learn the feature transform from raw data?

— Use the features to reconstruct the input with a decoder

— Autoencoder means the encoding itself



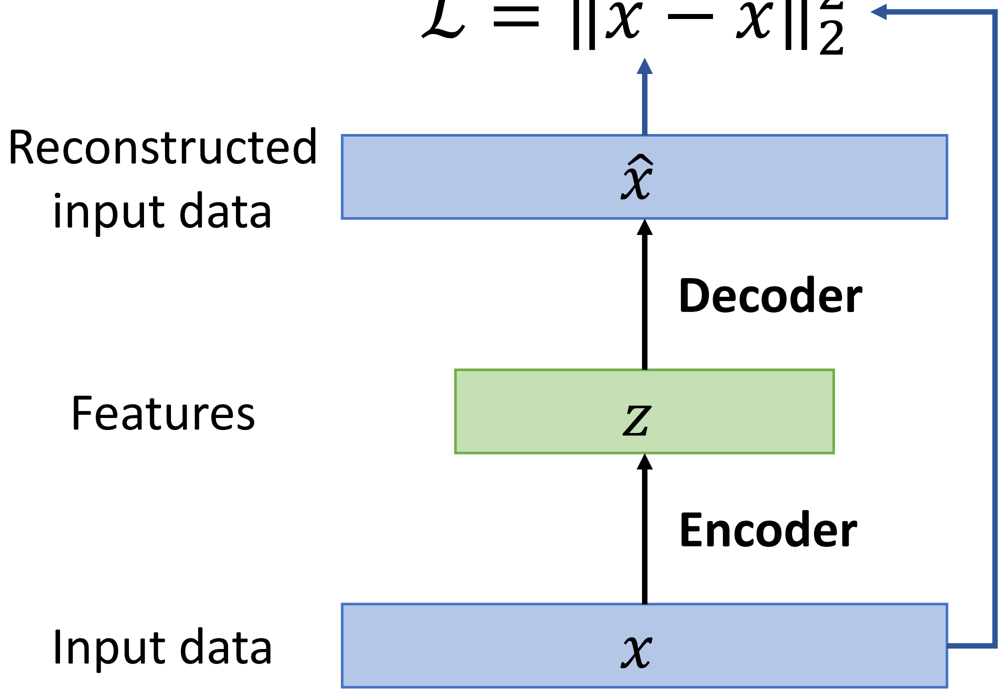
Auto-Encoders

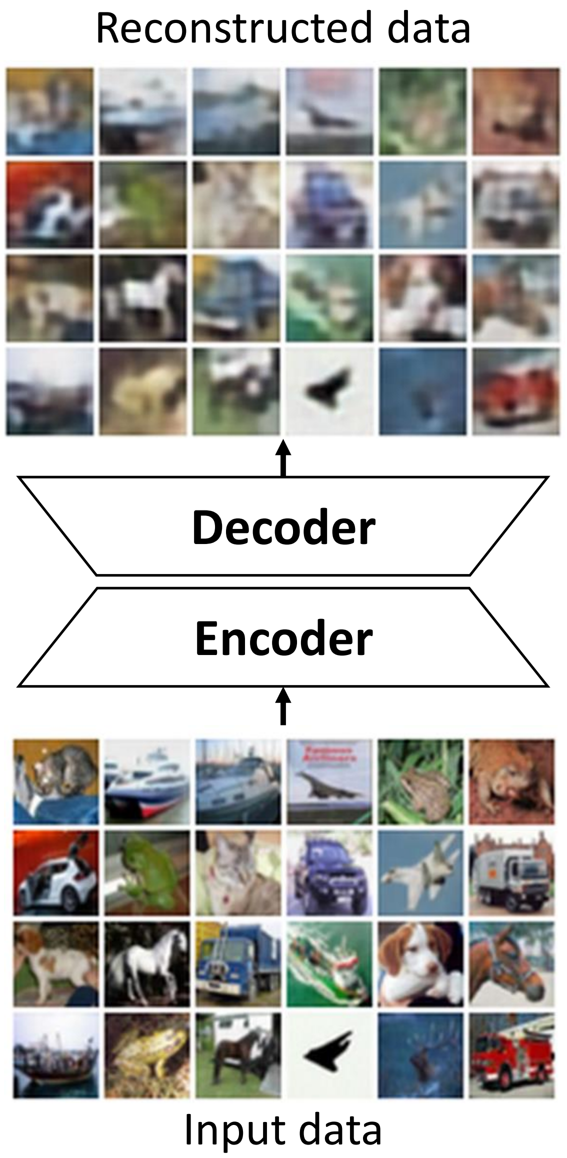
- How can we learn the feature transform from raw data?

— Use the features to reconstruct the input with a decoder

— Autoencoder means the encoding itself

$$\mathcal{L} = \|\hat{x} - x\|_2^2$$



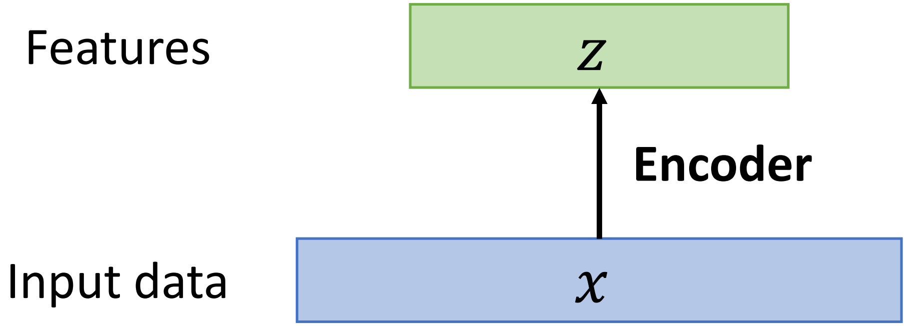


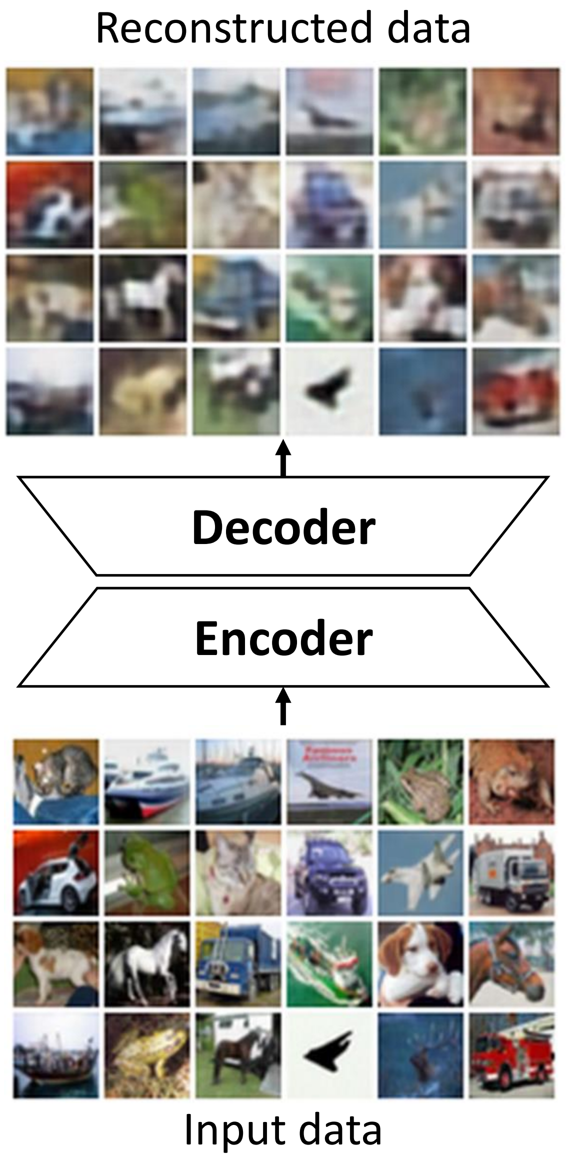
Auto-Encoders

- For a downstream task (out target application)

— Train autoencoders without labels

— After training, throw away the decoder





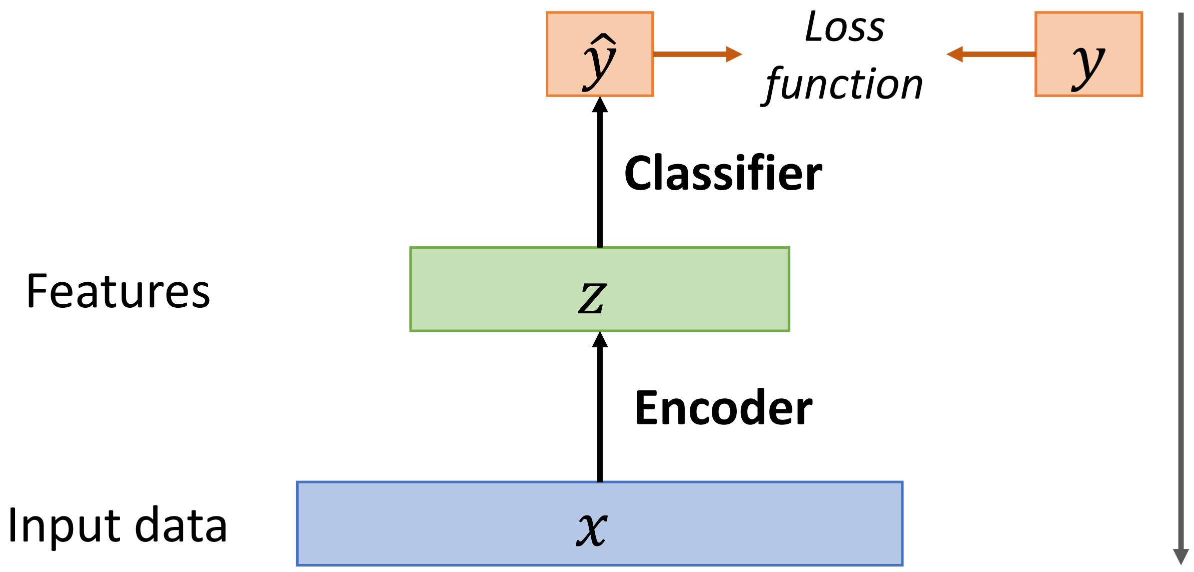
Auto-Encoders

- For a downstream task (out target application)

— Train autoencoders without labels

— After training, throw away the decoder

— Use the trained encoder only to extract latent features



Fine-tune the encoder
by training the classifier jointly

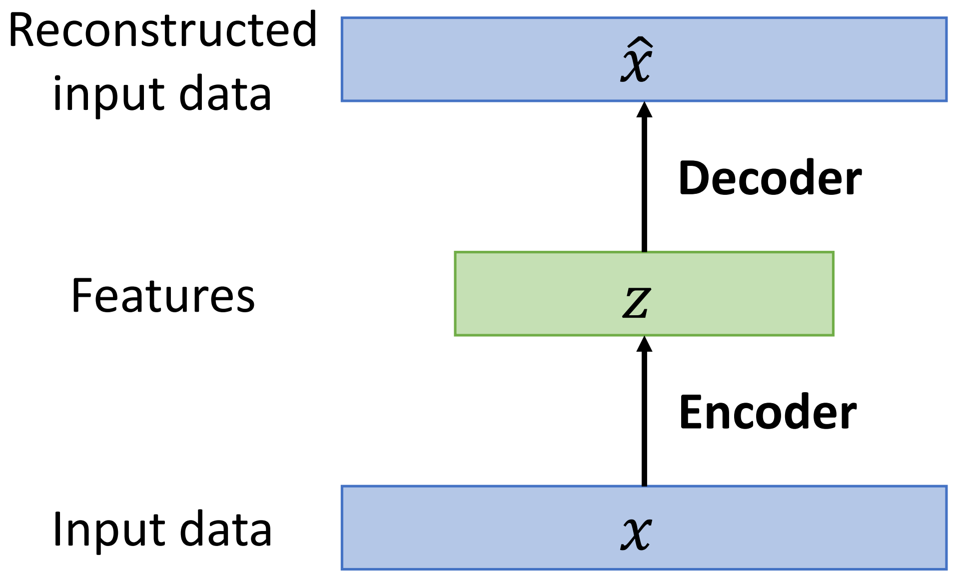
❖ Trained encoder can be used to initialize a supervised model

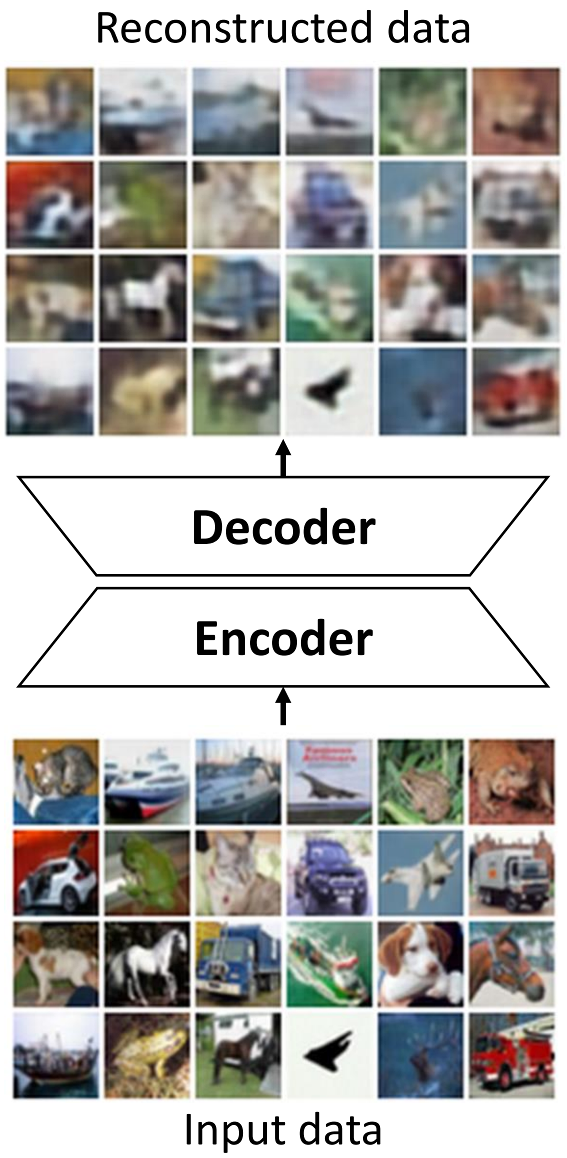
Auto-Encoders

- Autoencoders learn latent features without any labels

— It can use features to initialize a supervised model

— Not probabilistic: No way to sample new data from learned





Practice: Auto-Encoders (AE)

- Auto-Encoders
  - Image Reconstruction
  - De-noising