A Primer on Deep Learning
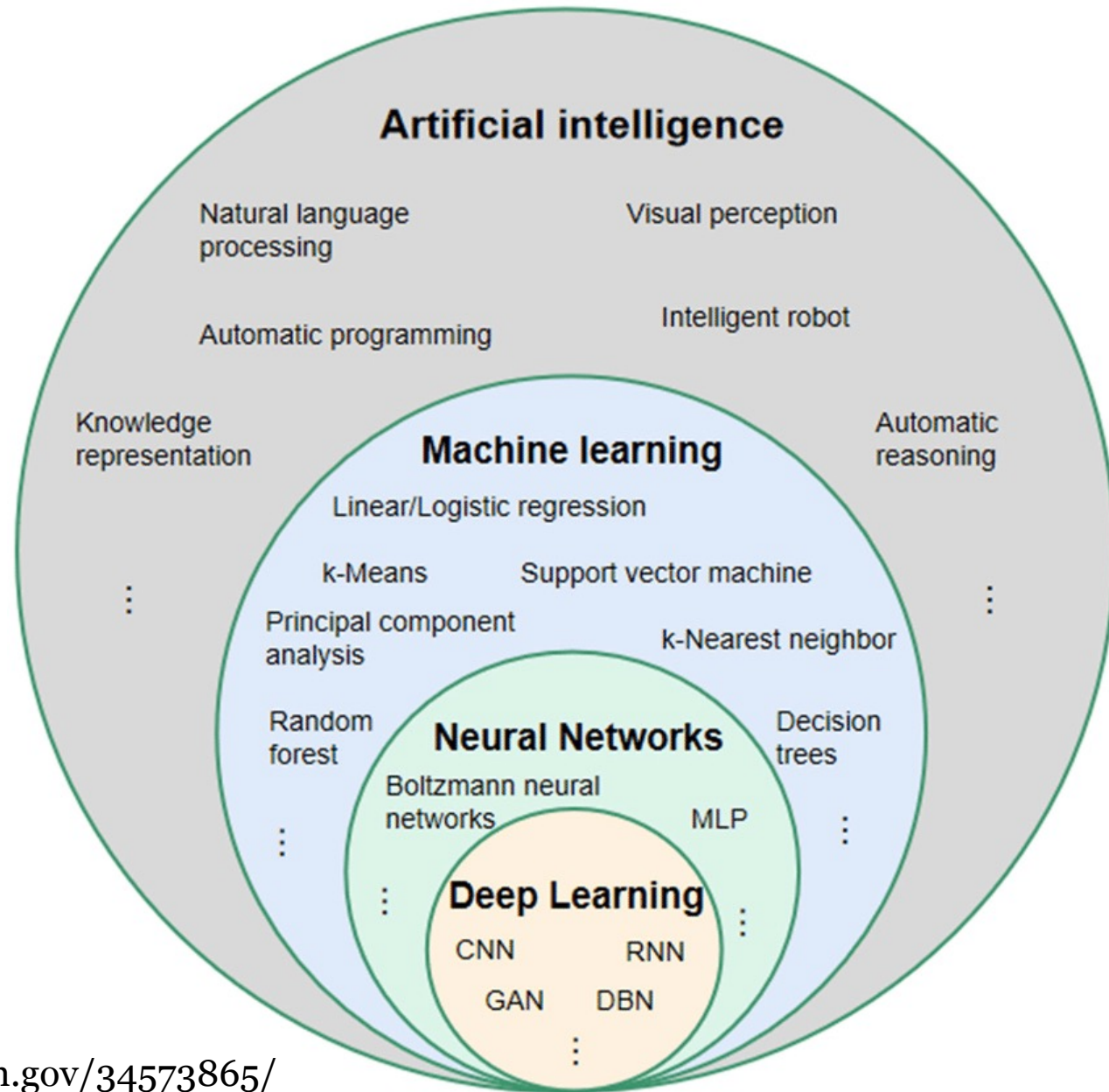
Dong Xu

EECS Department
C. S. Bond Life Sciences Center
University of Missouri, Columbia
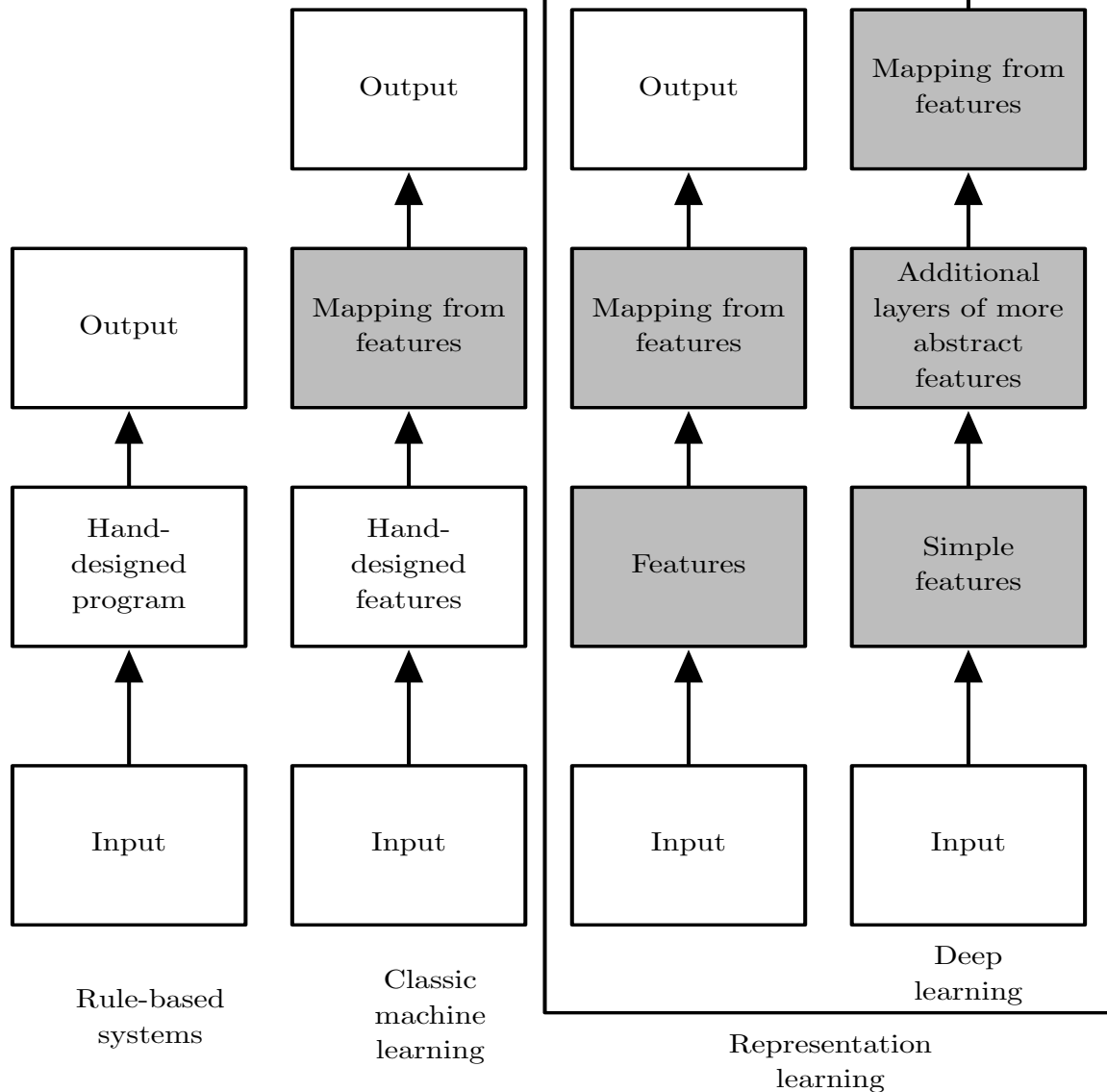*http://digbio.missouri.edu*

AI Scope

Different Approaches

Output

Output

Output

Mapping from
features

Output

Mapping from
features

Mapping from
features

Hand-
designed
program

Hand-
designed
features

Features

Additional
layers of more
abstract
features

Input

Input

Input

Simple
features

Input

Rule-based
systems

Classic
machine
learning

Deep
learning

Representation
learning

Supervised Machine learning

- Apply a prediction function to a feature representation of the image to get the desired output:

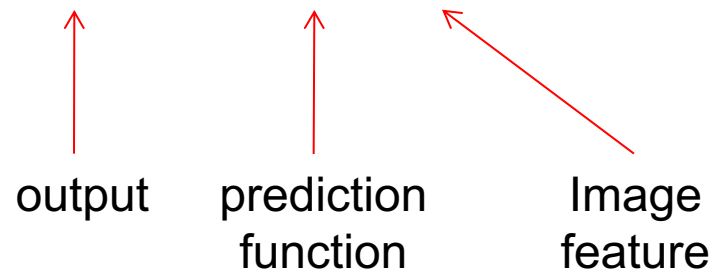$$f(\text{}) = \text{``apple''}$$

$$f(\text{}) = \text{``tomato''}$$

$$f(\text{}) = \text{``cow''}$$

Supervised Machine learning

$$y = f(\mathbf{x})$$

output   prediction function   Image feature

- **Training:** given a *training set* of labeled examples $\{(\mathbf{x}_1,y_1), ..., (\mathbf{x}_N,y_N)\}$, estimate the prediction function f by minimizing the prediction error on the training set

- **Testing:** apply f to a never before seen *test example* $\mathbf{x}$ and output the predicted value $y = f(\mathbf{x})$
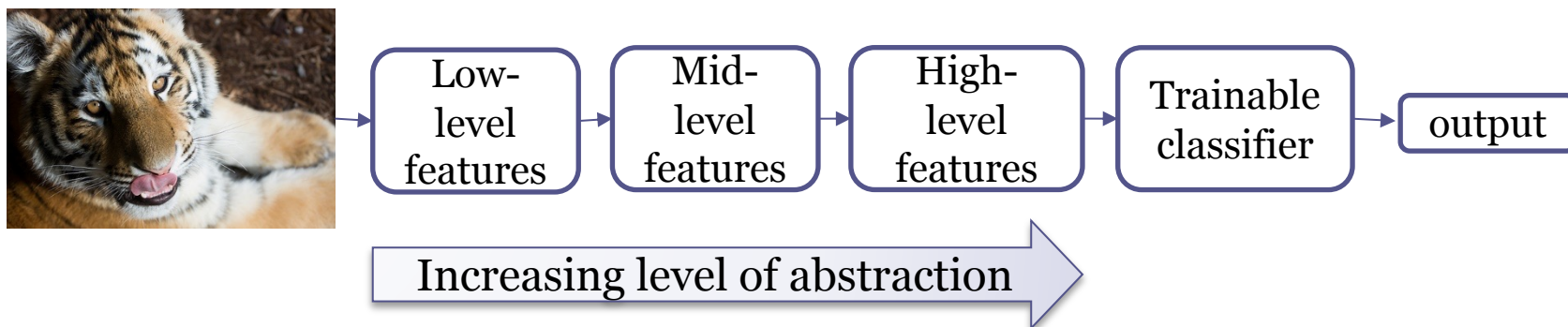
Generalization

Training set (labels known)

Test set (labels unknown)

- How well does a learned model generalize from the data it was trained on to a new test set?

Learning Hierarchical Representations

Increasing level of abstraction

- Hierarchy of representations with increasing level of abstraction. Each stage is a kind of trainable nonlinear feature transform

- Image recognition

  Pixel → edge → texton → motif → part → object

- Text

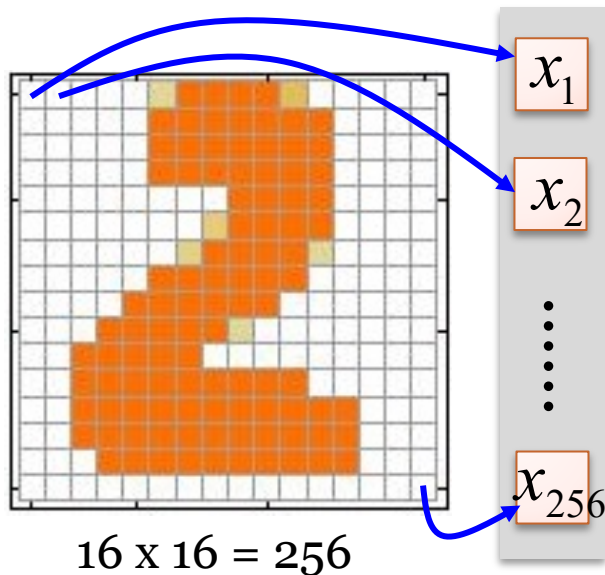 Character → word → word group → clause → sentence → story

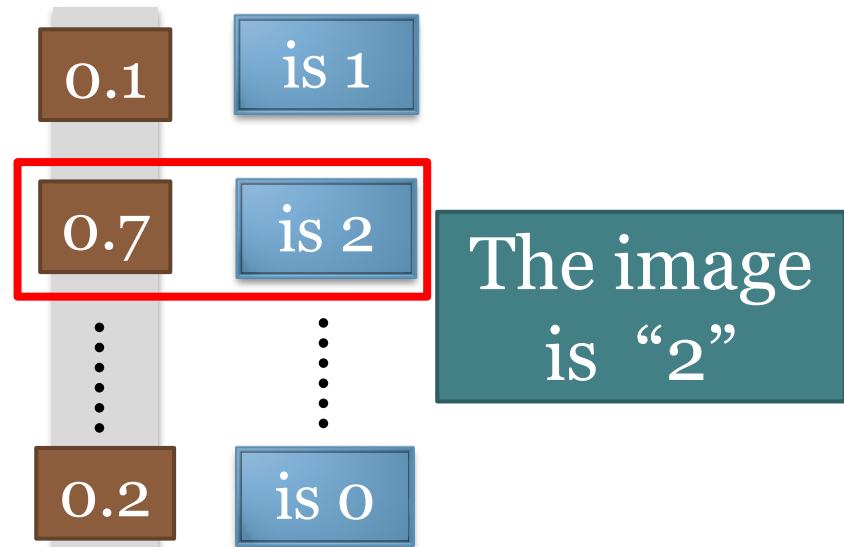# **Example Application**

- Handwriting Digit Recognition

$$f: R^{256} \rightarrow R^{10}$$
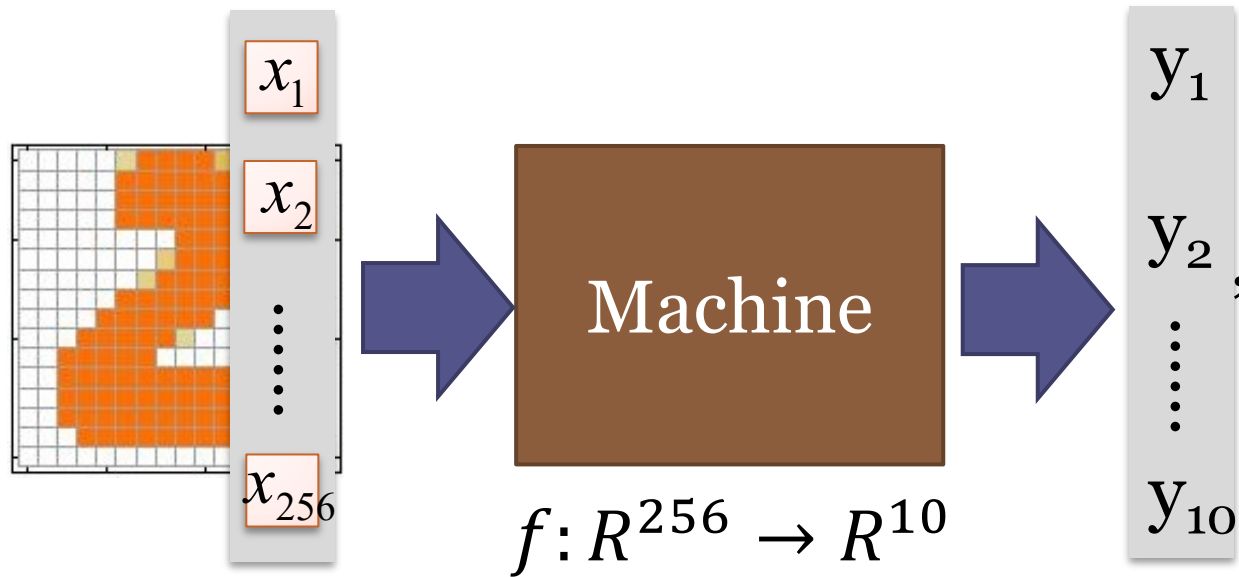
Handwriting Digit Recognition

Input

$16 \times 16 = 256$

Ink $\rightarrow 1$

No ink $\rightarrow 0$

Output

0.1	is 1
0.7	is 2
⋮	⋮
0.2	is 0

The image is "2"

Each dimension represents the confidence of a digit.

Deep Neural Network (DNN)
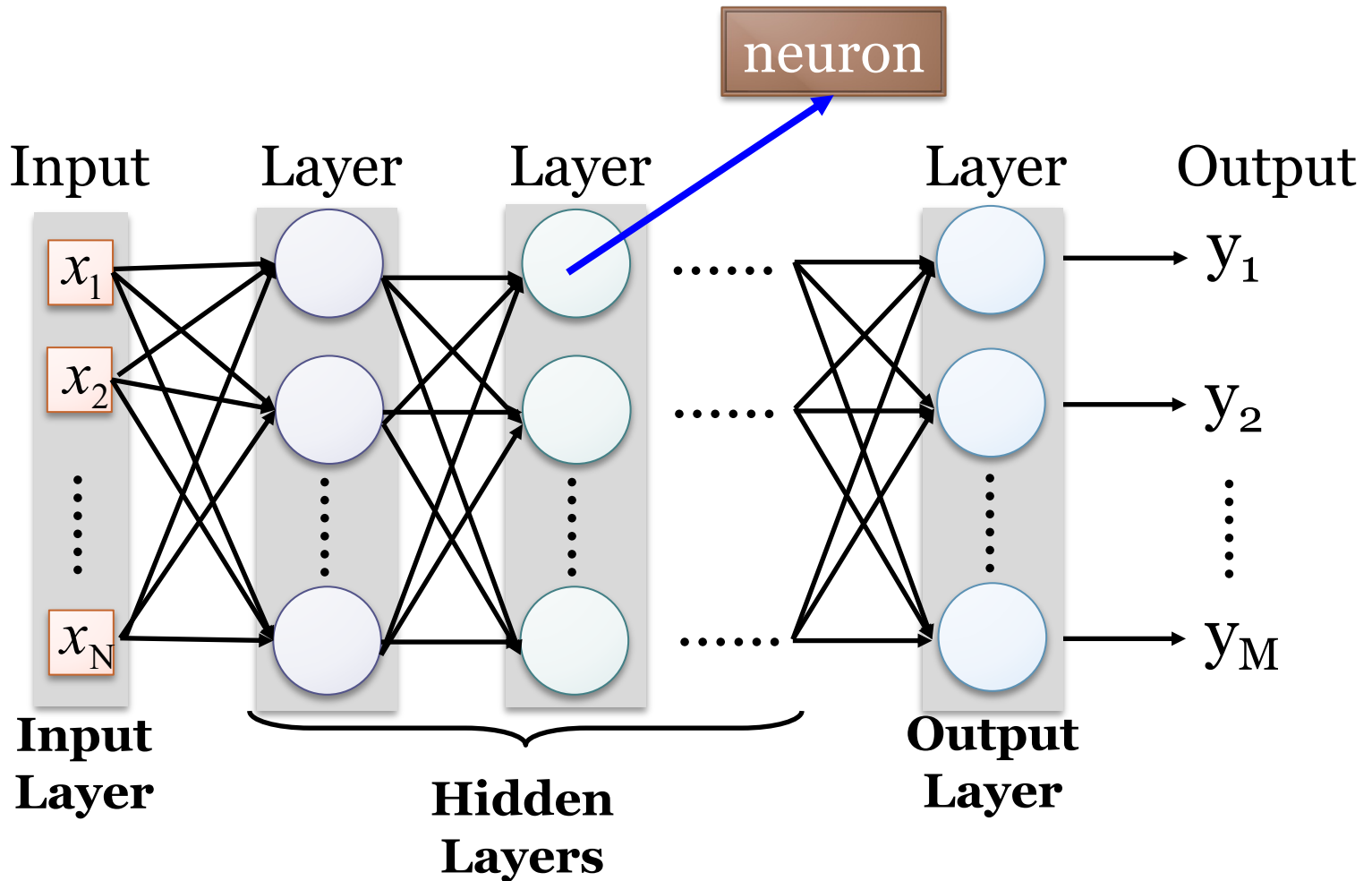
Abstraction and Representation

physical space → latent/embedding space (manifold)

photo impressionism expressionism cubism abstract expressionism
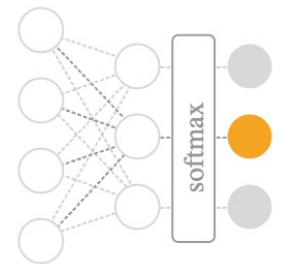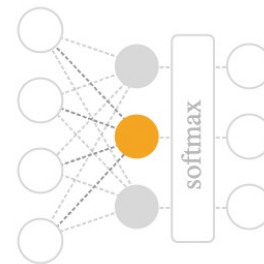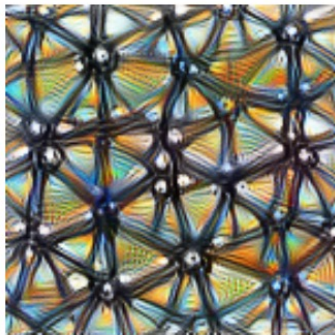
Latent Representation

Neuron	Channel	Layer/DeepDream	Class Logits	Class Probability
$layer_n[x,y,z]$	$layer_n[:,:,z]$	$layer_n[:,:,:]^2$	$pre\_softmax[k]$	$softmax[k]$

https://distill.pub/2017/feature-visualization/
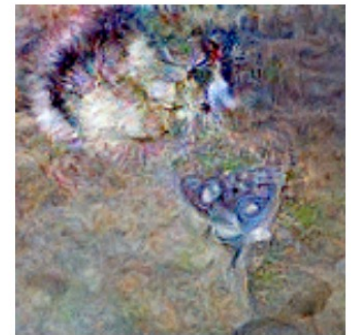
Unsupervised Learning & Embedding

High-dimensional signals
in a signal space S

Embedding space
(e.g., low-dimension,
small number of bits)
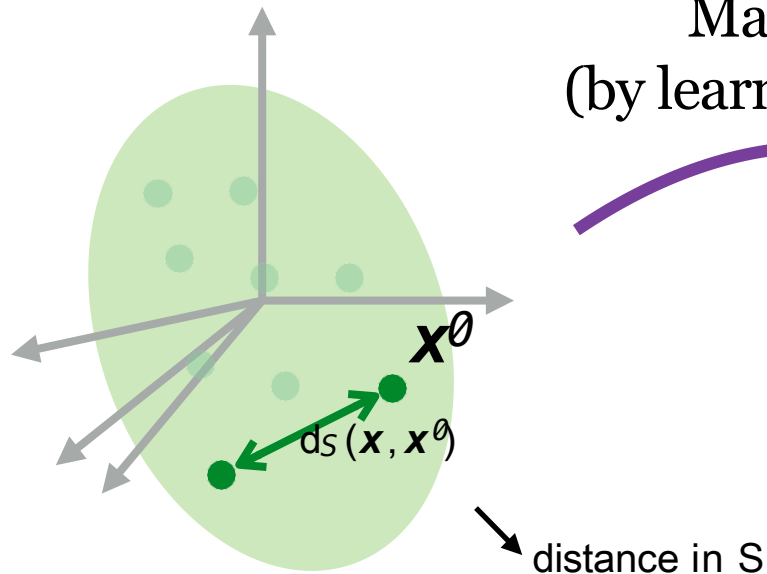
Mapping
(by learning in DL)

X^0

$d_S(x, x^0)$

distance in S
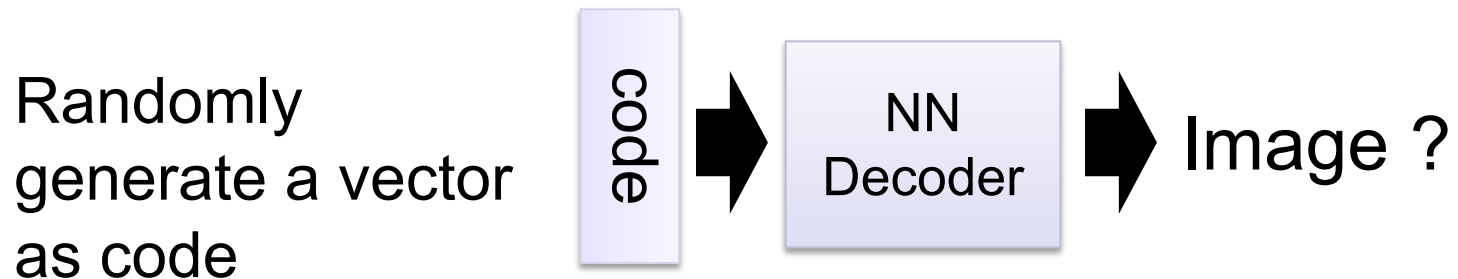
$f(x^0)$

$f(x)$

$d_E(f(x), f(x^0))$

distance in E
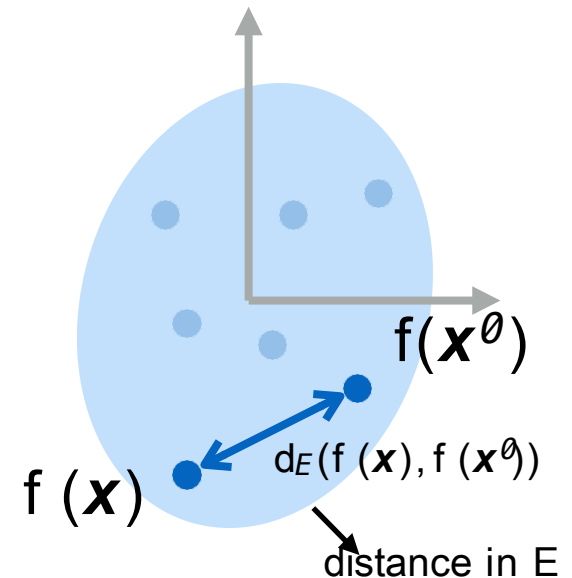
An embedding is a function from an original space to an embedding space that preserves aspects of the geometry of the original space

Autoencoder

As close as possible
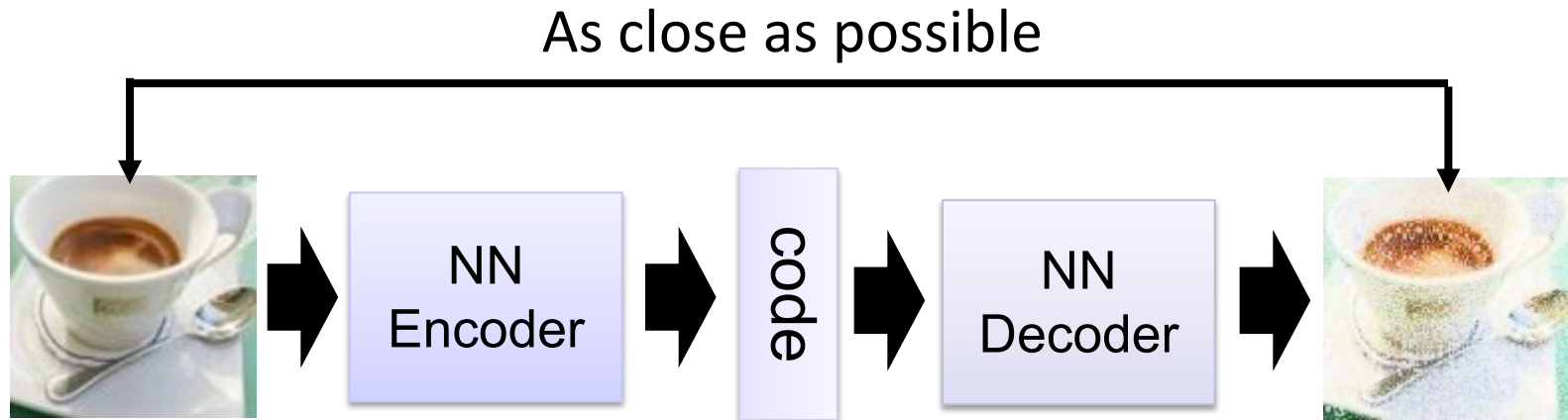
A type of unsupervised learning which discovers generic features of the data (**learn data patterns**)

Randomly generate a vector as code

Autoencoder with 3 Layers
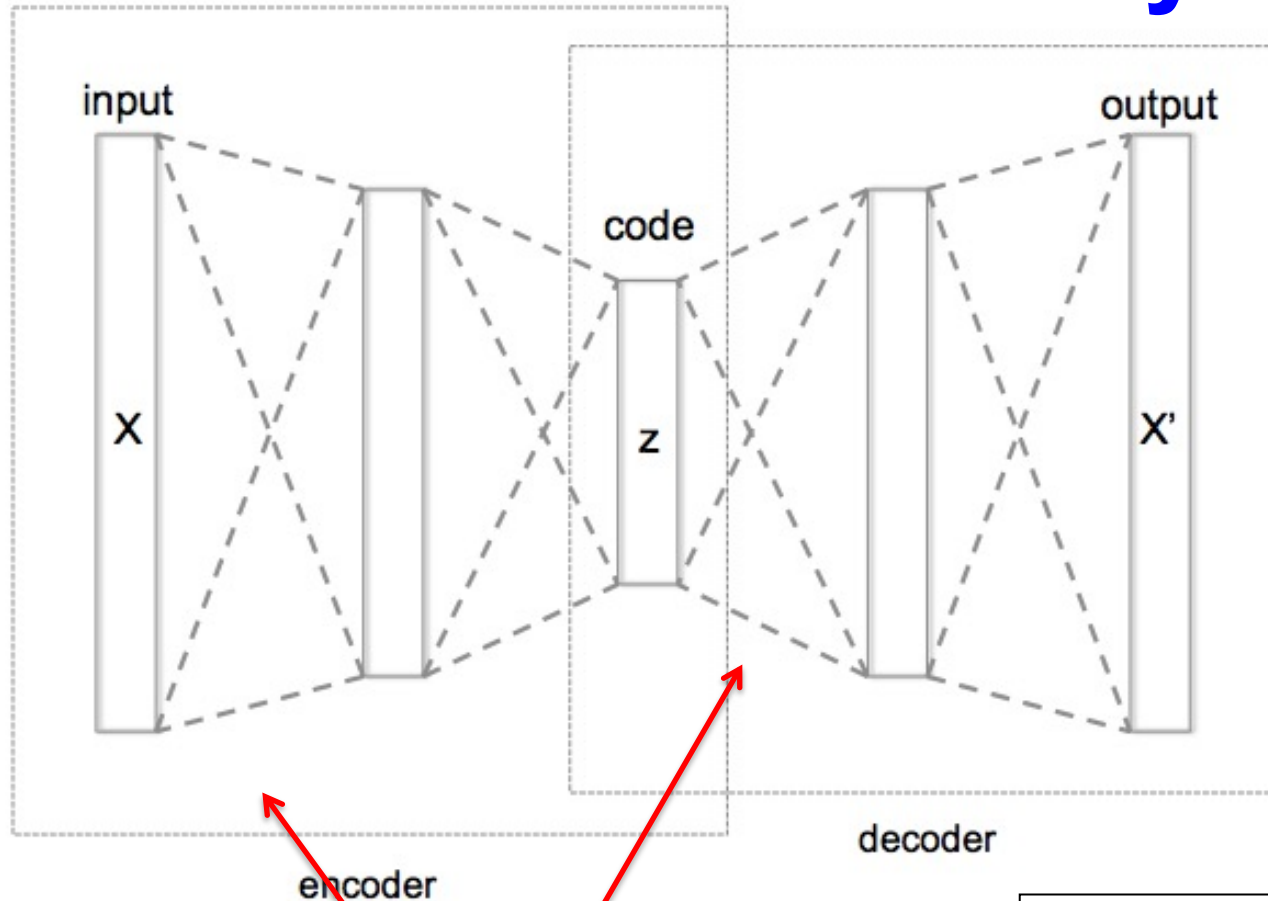
Large → small, learn to compress

Z: Latent variable; embedding

Training: model.fit(X,X)
Cost function: $\Sigma_{k=1..N} (x_k - x'_k)^2$

T-SNE Visualization

t-Distributed Stochastic Neighbor Embedding (**t-SNE**)

minimizes the KL divergence between distributions

U-Map

Uniform Manifold Approximation and Projection (UMAP)

find a topological representation of the data in a lower dimensional space through manifold learning technique

https://github.com/lmcinnes/umap
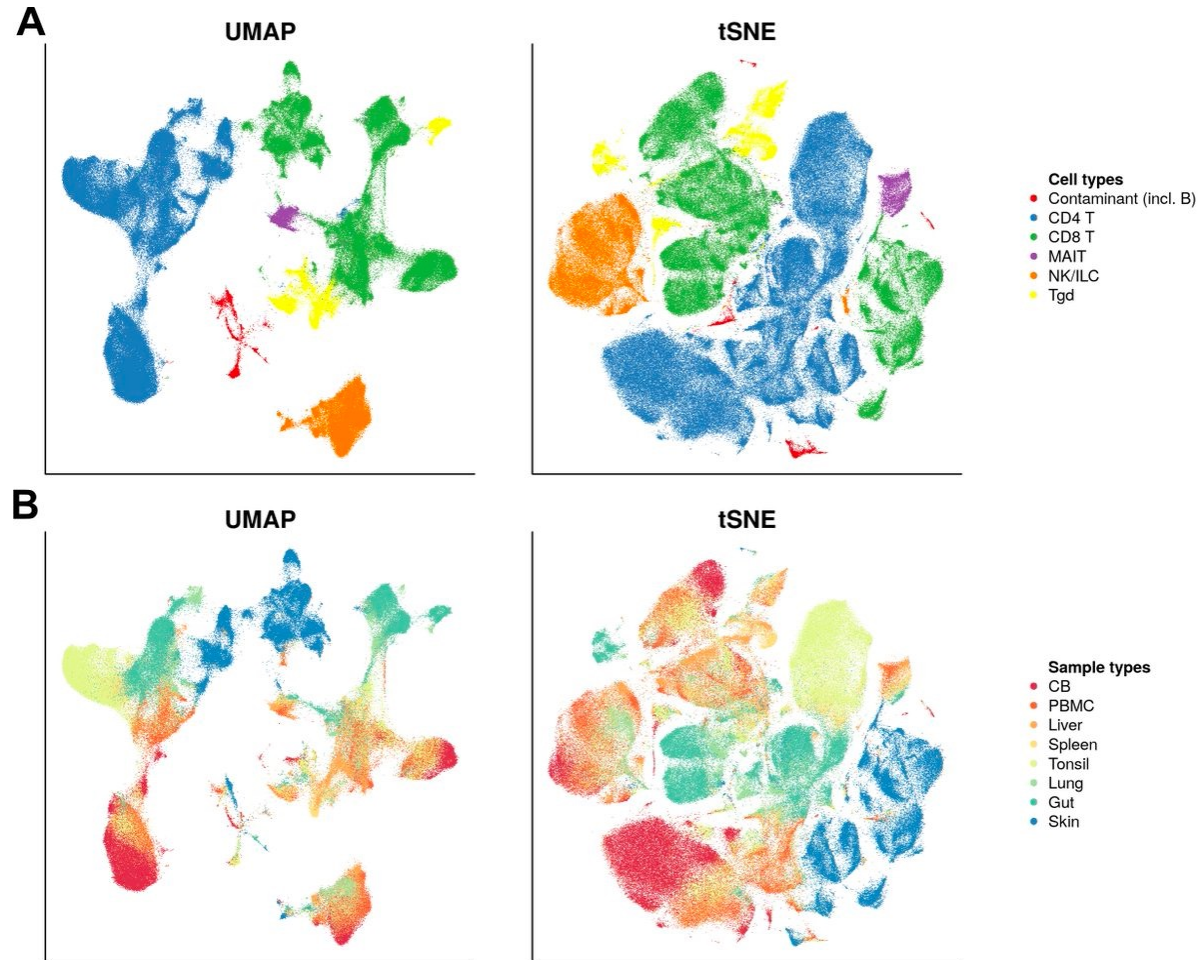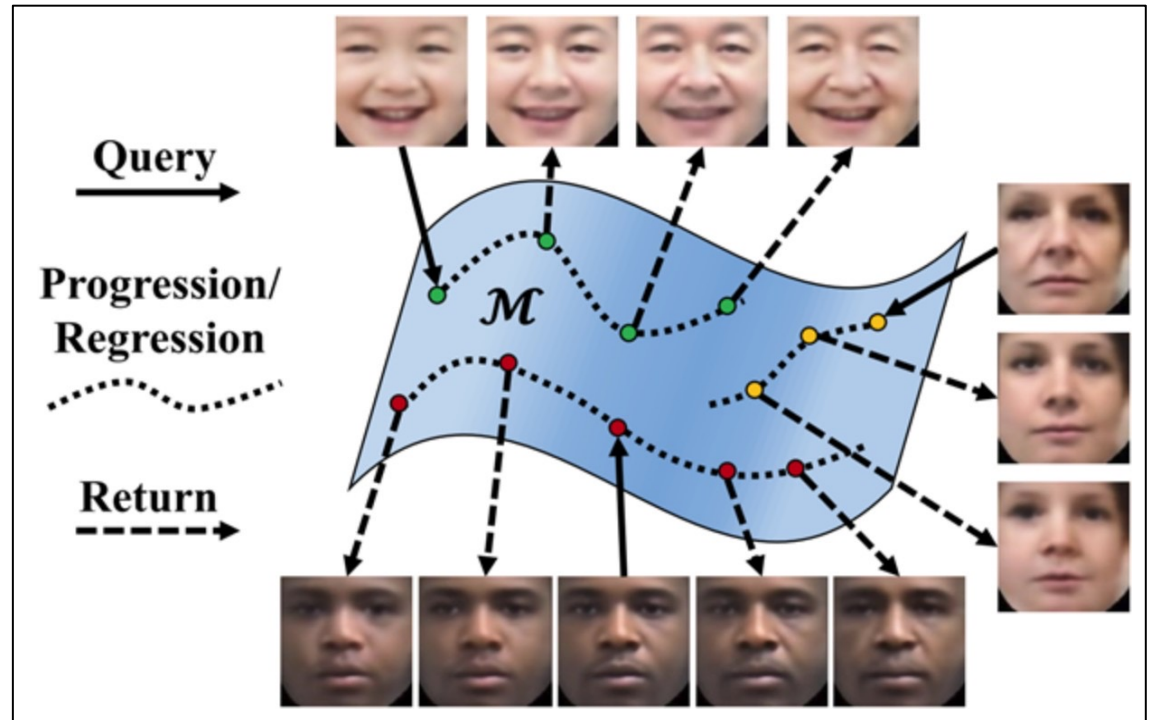
Manifold

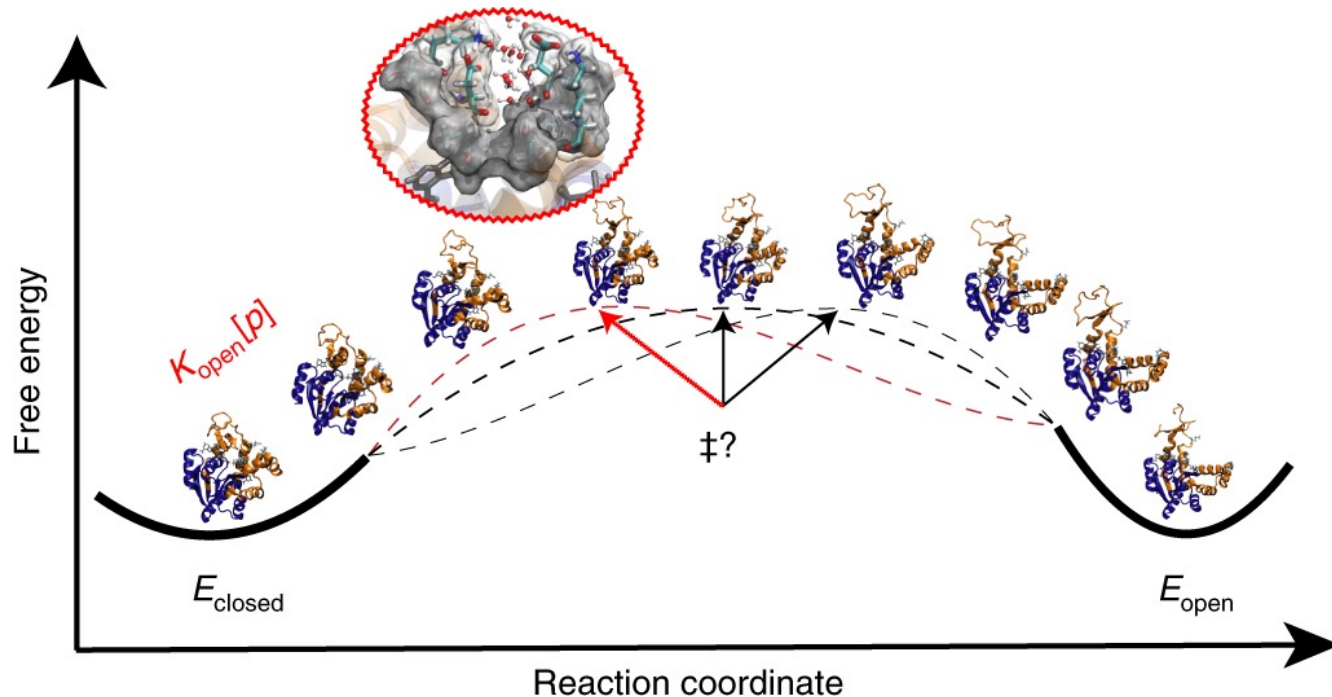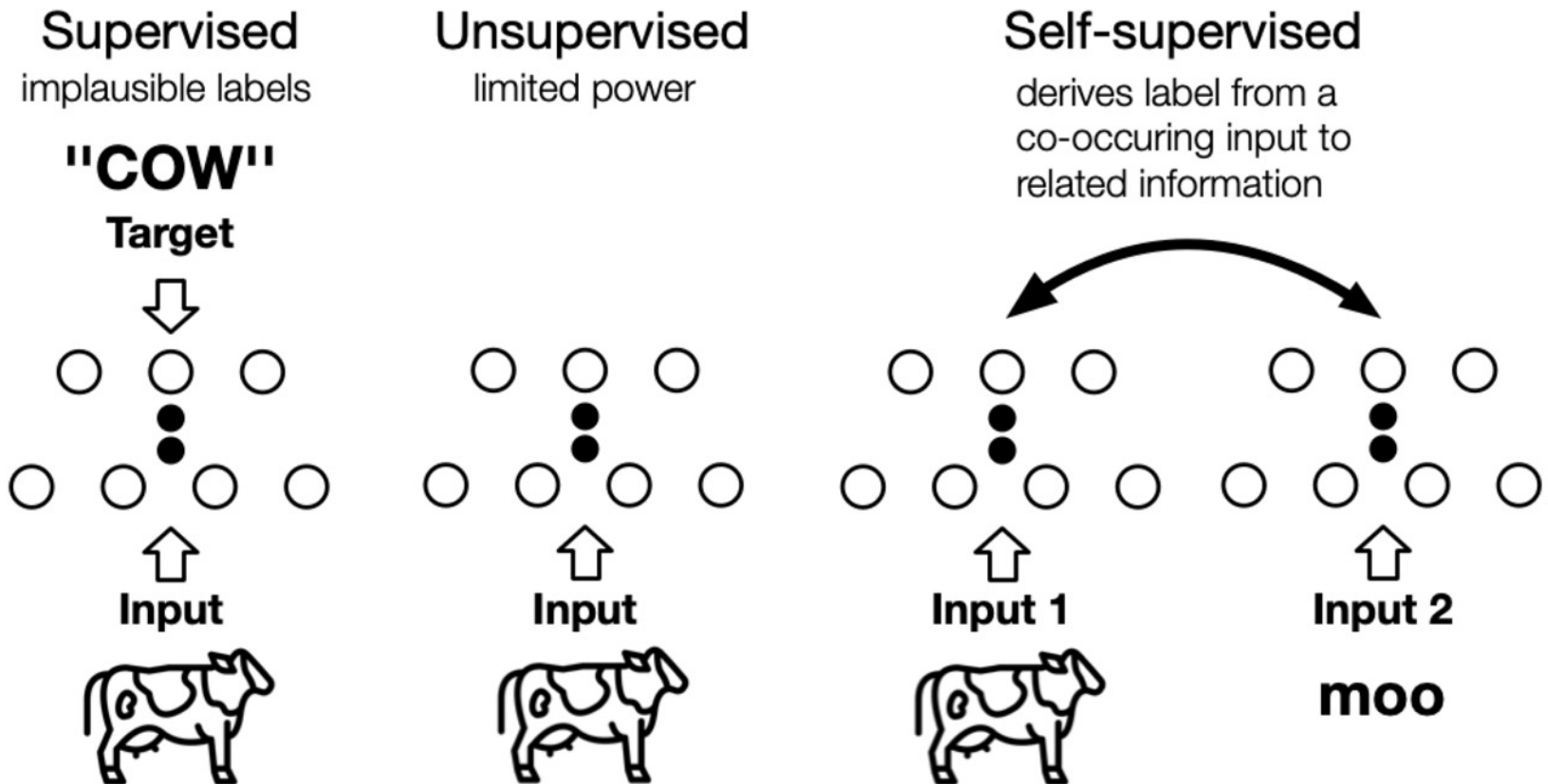- A Manifold is a topological space that locally resembles Euclidean space near each point
- n-dim manifold → topological space M, every point x ∈ M has a neighbor homeomorphic (isomorphismic) to Euclidean space Rn

Manifold Hypothesis

- **DL Central Hypothesis**: Data concentrates around a low-dim manifold (**relevant dimension**)
- Mimic human learning
- Not all embeddings produce manifold

Self-supervised learning

Supervised
implausible labels

"COW"

Target

⇩

○ ○ ○
●
●
○ ○ ○ ○ ○

⇧

Input

Unsupervised
limited power

○ ○ ○
●
●
○ ○ ○ ○ ○

⇧

Input

Self-supervised

derives label from a
co-occuring input to
related information

○ ○ ○   ○ ○ ○
●   ●
●   ●
○ ○ ○ ○ ○   ○ ○ ○ ○ ○

⇧   ⇧

Input 1   **Input 2**

moo

representation learning that enables learning good data representation from unlabeled dataset.

Foundation model era

- Machine learning paradigms

  Feature engineering: manual feature extraction (SVM, LightGBM, XGBoost)

  Architecture engineering: raw features, design deep network (CNN, LSTM)

  Objective engineering: pre-train large model and fine-tune it (ResNet50, Bert)

  Prompt engineering: prompt **foundation model** in zero/few shots

- Industrial era of artificial intelligence

AM radio receiver circuit

ElecCircuit.com