SEMINAR IN DATA SCIENCE AND INFORMATION TECHNOLOGY

ADVANCED LEARNING REPRESENTATIONS

LECTURE 1: OVERVIEW

WEEK 1-11

Instructor： Yang Li

SPRING 2023

LECTURE OUTLINE

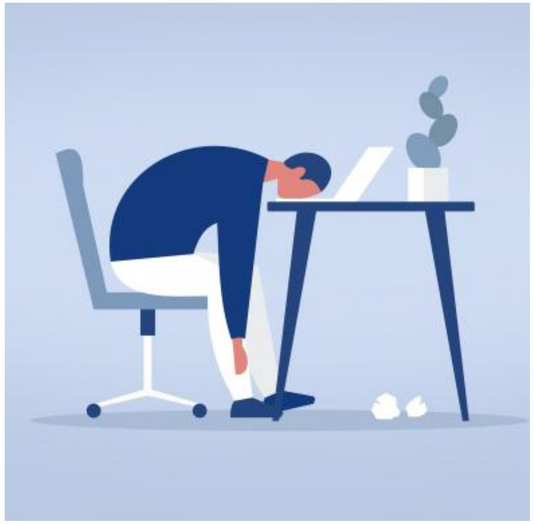
- About this class
- A preview of the topics we will cover
- Break
- Discussion on *Representation Learning: A New Perspective*

PURPOSE

- **What have we learned so far in Learning from Data:**
  - **Supervised Learning: discriminative & generative models, SVM, neural net**
  - **Unsupervised Learning: clustering & linear projection models (PCA/ICA/CCA)**
  - **Semi-Supervised Learning**
  - **Reinforcement Learning**

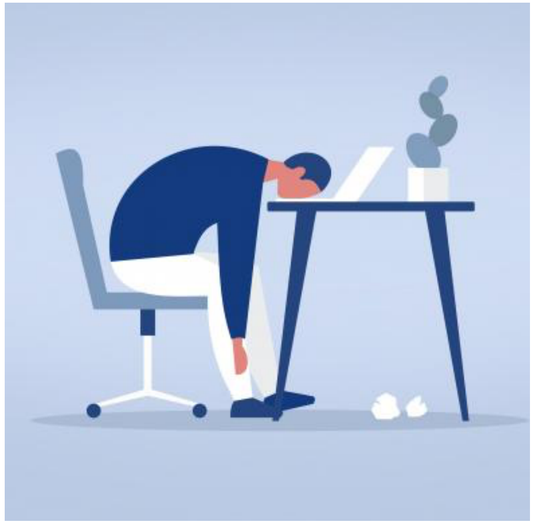
PURPOSE

- **What's next after Learning from Data?**



PURPOSE

- **What's next after Learning from Data?**

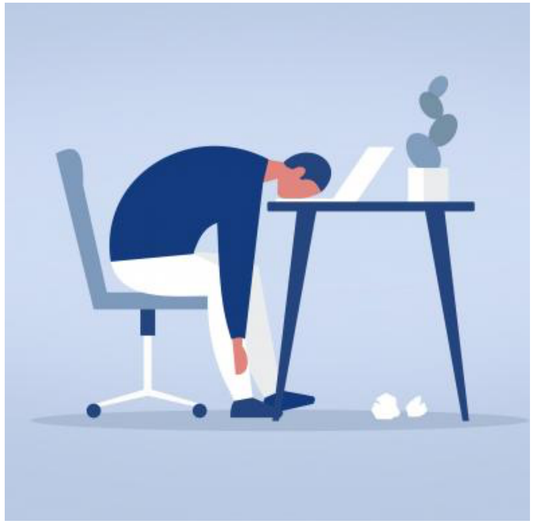


Given a new problem, how to find the best learning method?

PURPOSE

- **What's next after Learning from Data?**

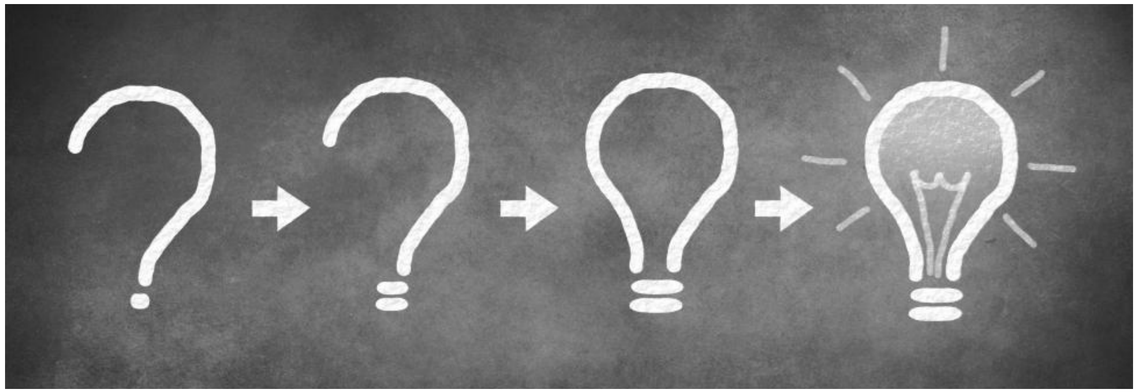
How to find "ideas" for machine learning research?



Given a new problem, how to find the best learning method?

PURPOSE

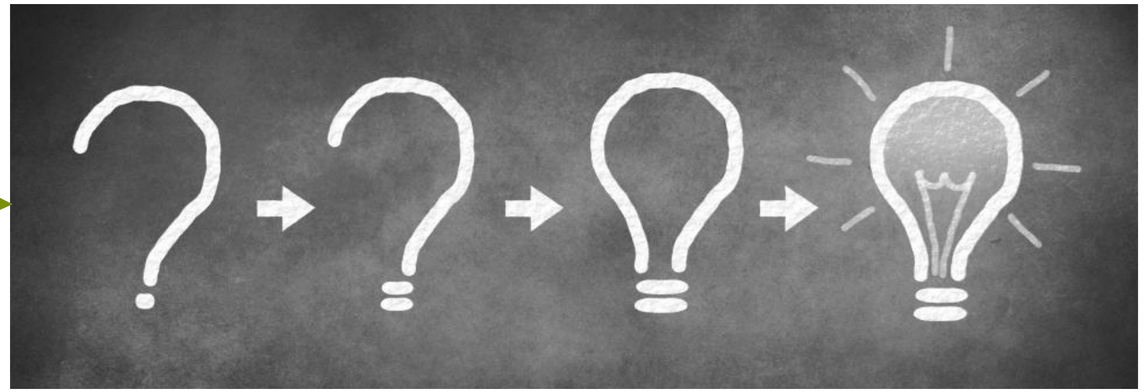
- **Learn from classic literatures**



PURPOSE

- **Learn from classic literatures**

Where did it draw inspiration from?

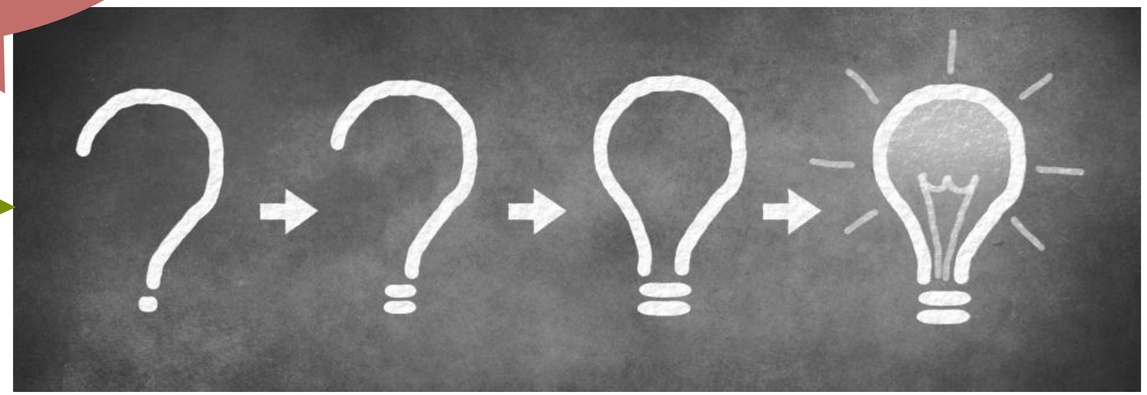


PURPOSE

- **Learn from classic literatures**

How was the problem formulated?

Where did it draw inspiration from?



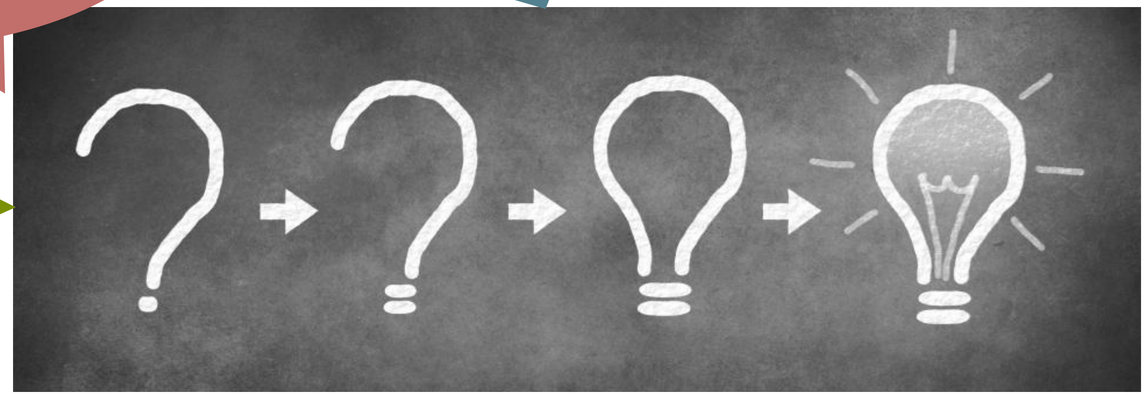
PURPOSE

- **Learn from classic literatures**

How was the problem formulated?

What priors or assumptions did it make?

Where did it draw inspiration from?



PURPOSE

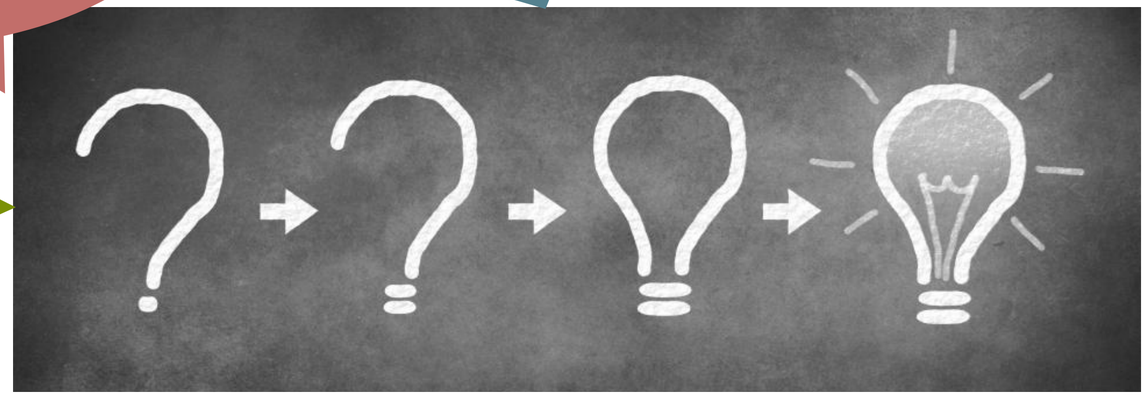
- **Learn from classic literatures**

How was the method derived?

What priors or assumptions did it make?

How was the problem formulated?

Where did it draw inspiration from?



PURPOSE

- **Learn from classic literatures**

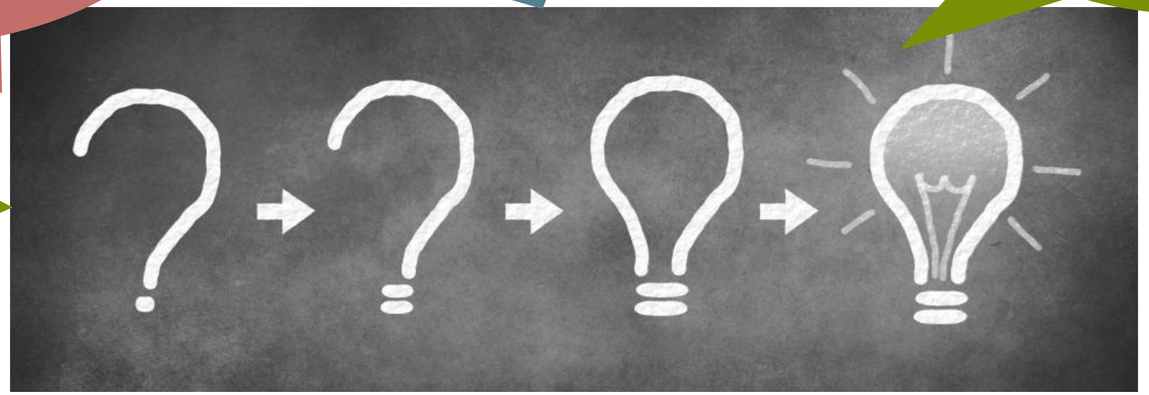
How was the problem formulated?

What priors or assumptions did it make?

How was the method derived?

How did it contribute to the field?

Where did it draw inspiration from?



PURPOSE

- **Learn from classic literatures**

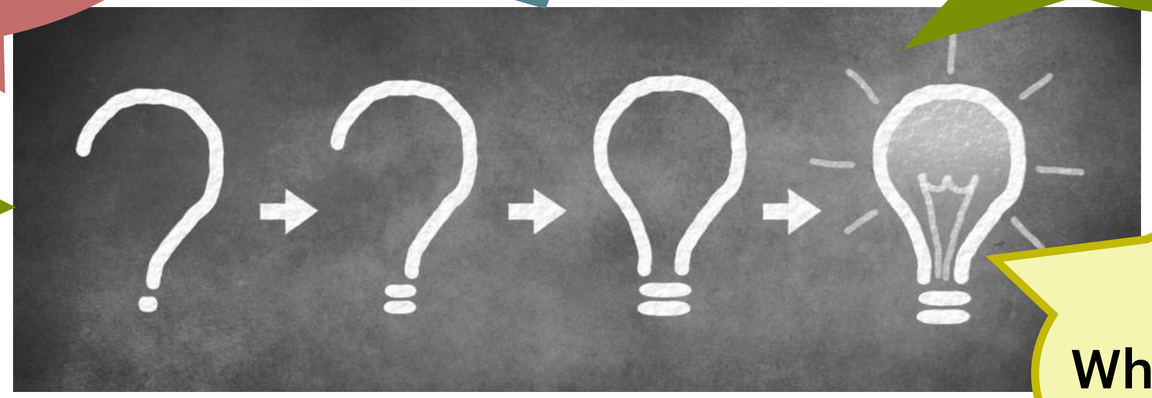
How was the problem formulated?

What priors or assumptions did it make?

How was the method derived?

How did it contribute to the field?

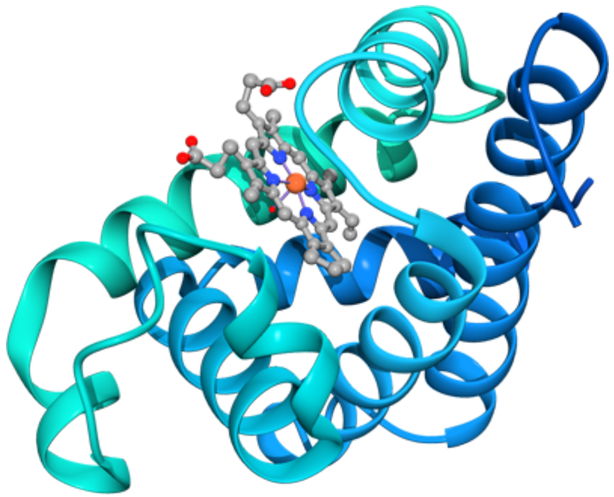
Where did it draw inspiration from?

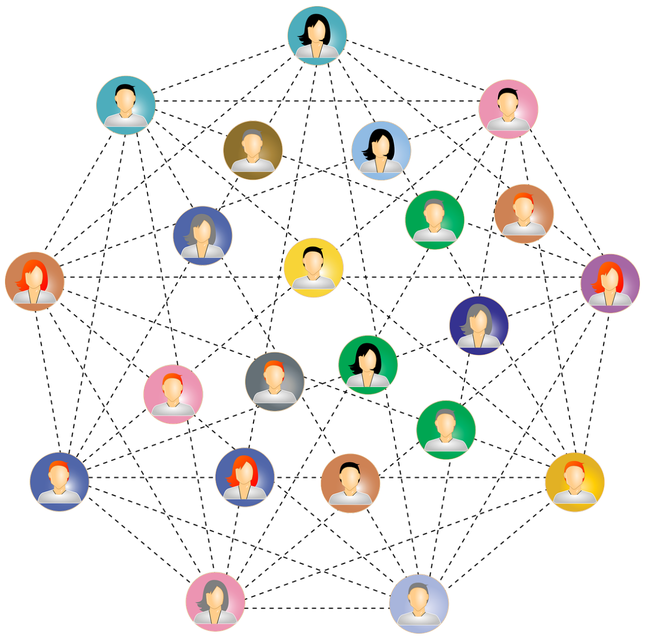


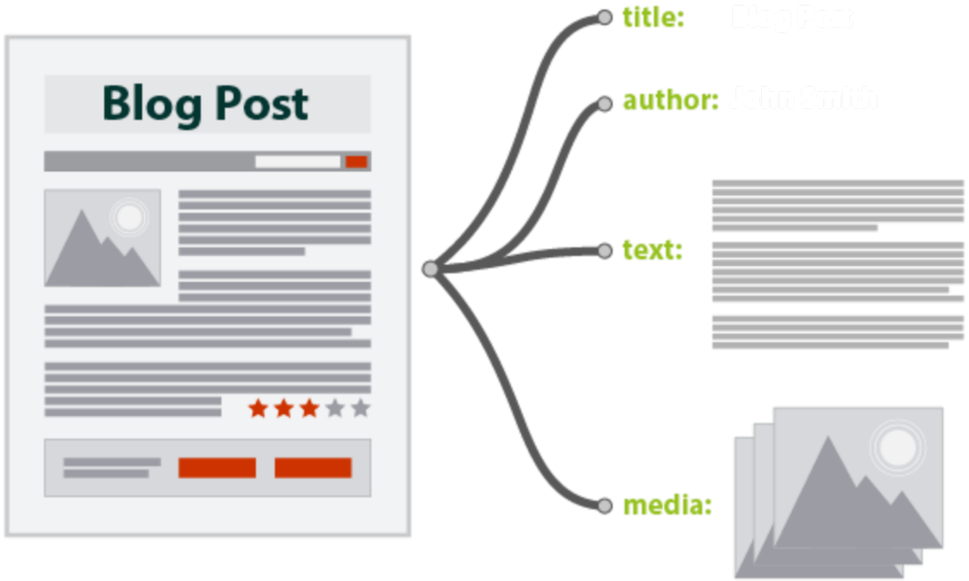
What can I learn from this work?

REPRESENTATION LEARNING

- **"Deep learning is about learning good representations"**



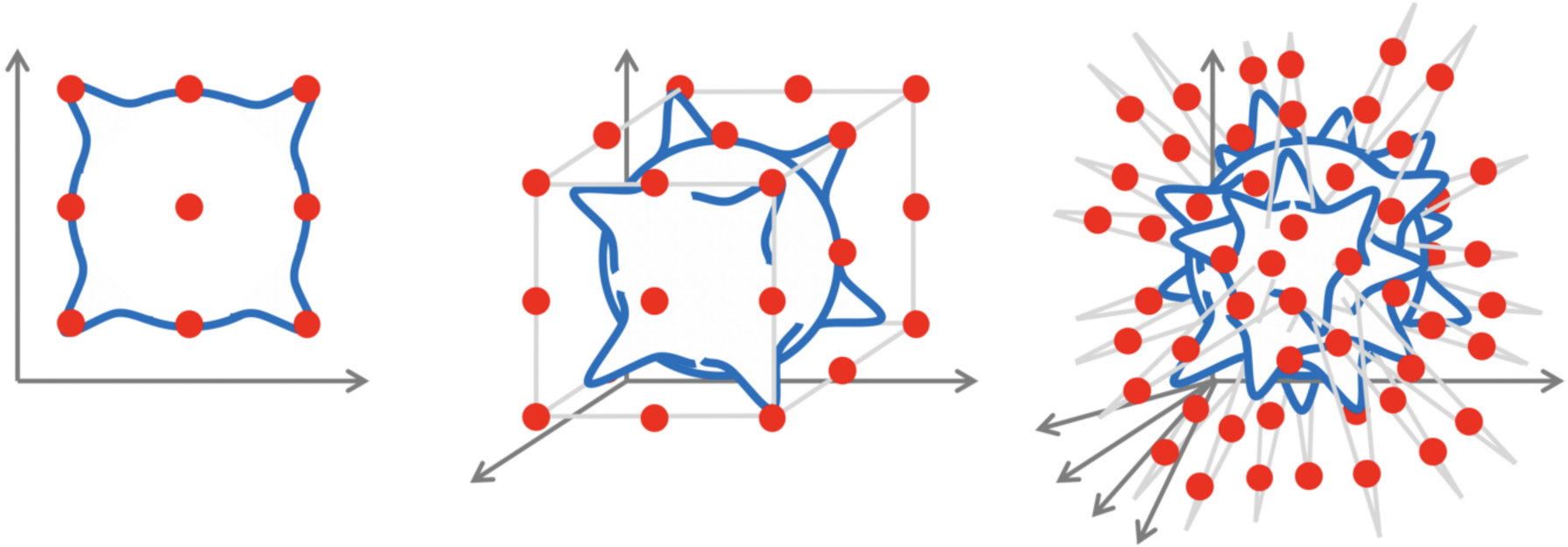




Examples of high dimensional, structured data

REPRESENTATION LEARNING

- **"Curse of dimensionality"**



Approximating a Lipschitz-continous function in R^d within error ε requires at least $\mathcal{O}(1/\varepsilon^d)$ samples

REPRESENTATION LEARNING

- This class focus on general paradigms for learning good representations
  - How to represent structured data in learning ? e.g. images, text, sequences, graphs
  - How to exploit the hidden structures and prior knowledge in data?

Learn from classic literature

SEMINAR FORMAT

- Read and discuss classic literature each week
  - Student will rotate to present the reading material
  - Form groups of 3-4 people for in-class discussion
- In-class programming exercises
- *Guest lecture

HOMEWORK

- Submit before each class
  - Write a 1-page reading summary
  - Answer pre-discussion questions

PRESENTATIONS & FINAL REPORT

- Each group needs to do at least one presentation
  - One of the required readings (Week 2-9)
  - A topic related to the required reading (Week 10 & 11)
- Group report:
  - A written report based on the topic your presented

GRADING

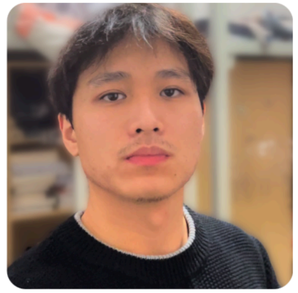
- **Homework: 60%**
- **Class participation: 20%**
- **Final presentation & report: 20%**

CLASS LOGISTICS

- **Join class WeChat group**
- **Class website**
- **Contact TA if you need to use the class GPU server**

http://yangli-feasibility.com/home/classes/dss2023spring/

Yang Li
Instructor
yangli@sz.tsinghua.edu.cn



Tao Shi
TA
shitao21@mails.tsinghua.edu.cn

GuanLin Wu
TA
18342291462@163.com

Office hours

Name	Time	Location
Yang	Friday 2:00-4:00pm	Info Building 1108a
Tao	Monday 2:00-4:00pm	Info Building, 1111A
Guanlin	Monday 10:00-12:00am	Info Building, 702

You can also make appointments outside office hours.

LECTURE OUTLINE

- About this class
- A preview of the topics we will cover
- Break
- Discussion on *Representation Learning: A New Perspective*

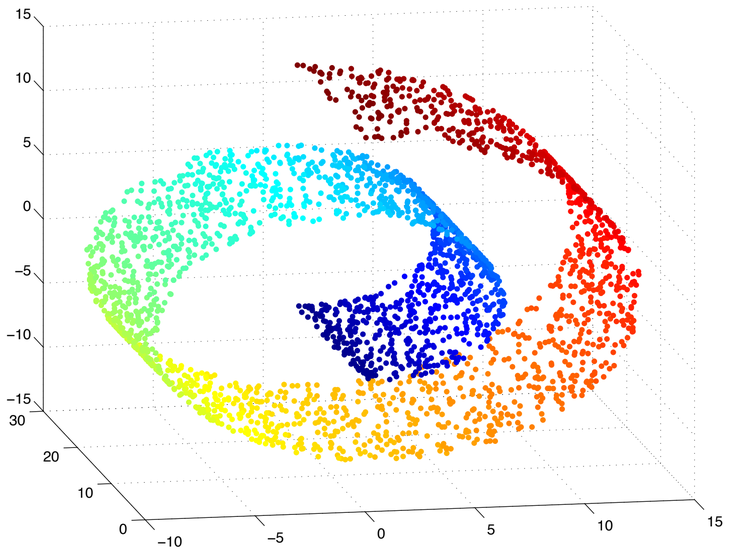
MAIN TOPICS

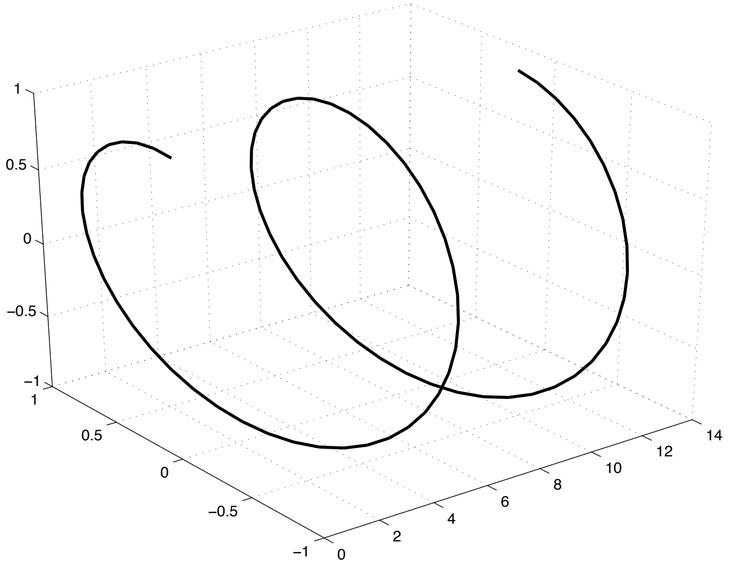
1. **Manifold learning**
2. **Generative modeling**
3. **Metric learning**
4. **Attention mechanism**
5. **Graph representation learning**

1. MANIFOLD LEARNING

- **Non-linear Dimension Reduction (NLDR): Recover meaningful low-dimensional structures hidden in high-dimensional data.**

- **Manifold Learning :Given *m* points $x^{(1)}, \ldots, x^{(m)}$ on a manifold M embedded in $\mathbb{R}^D$, find a set of points $y^{(1)}, \ldots, y^{(m)} \in \mathbb{R}^d, (d \ll D)$ such that $y^{(i)}$ represents $x^{(i)}$**

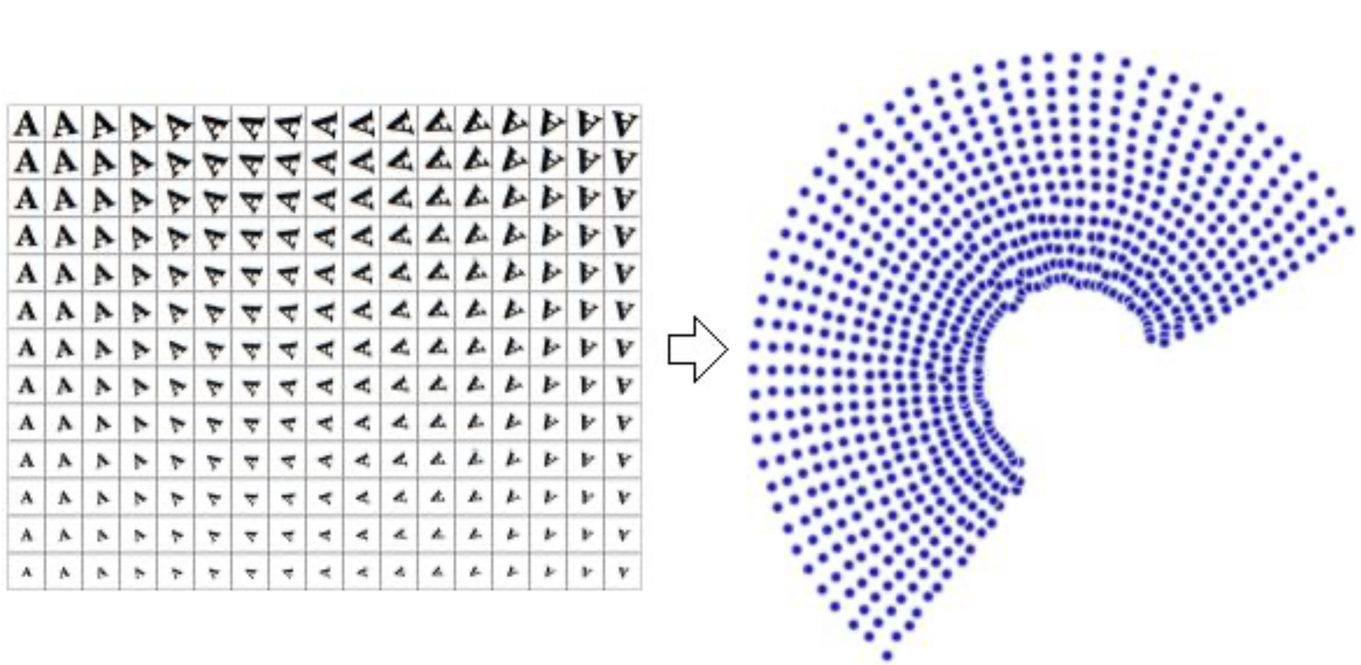


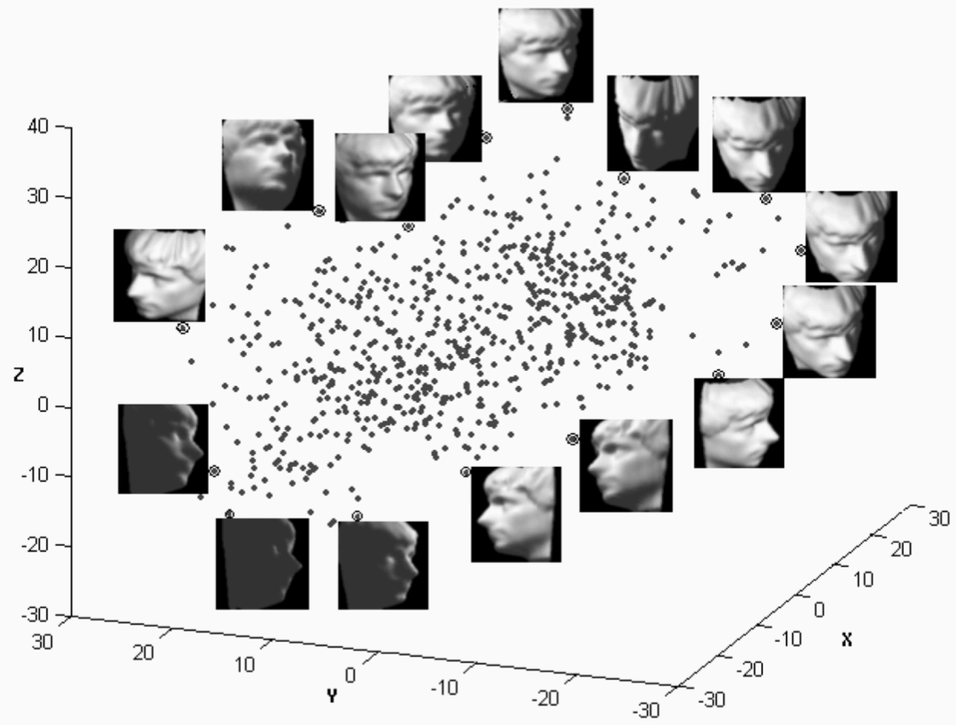


Manifold (informal): a latent space of a given dimension (degree of freedom)

1. MANIFOLD LEARNING

Examples of data manifold: *image dimension reduction*

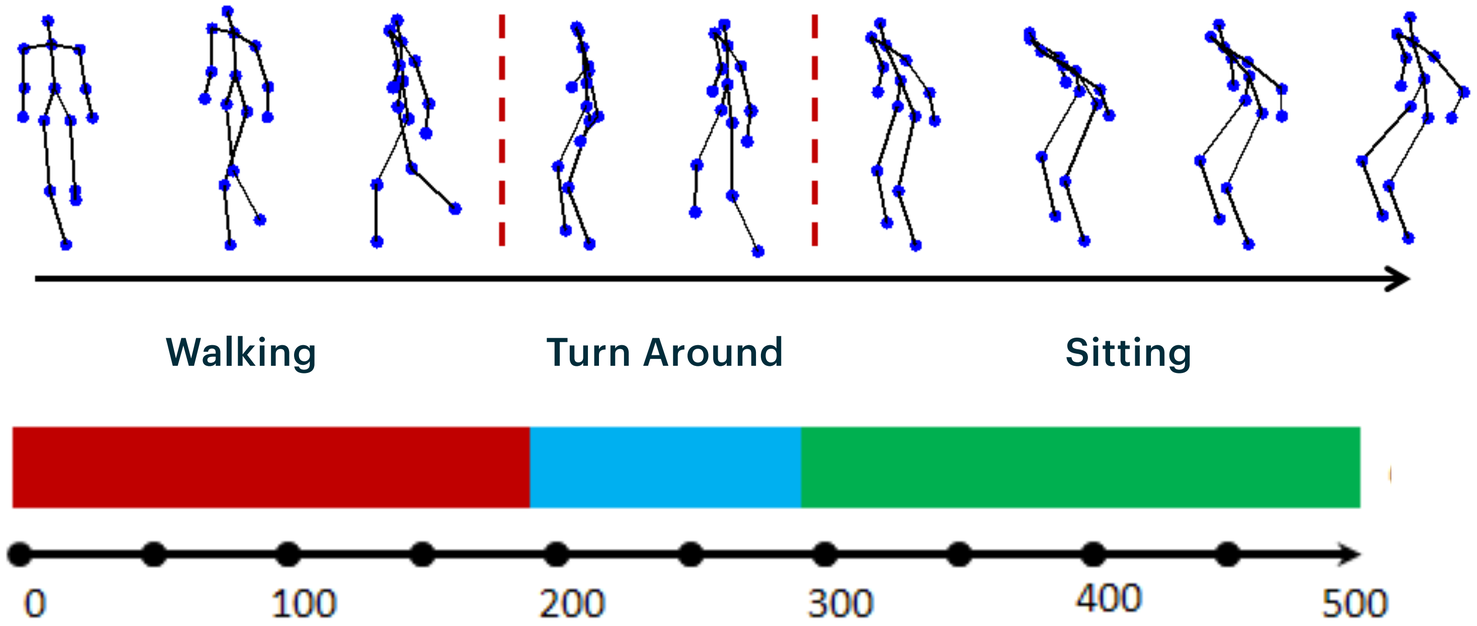


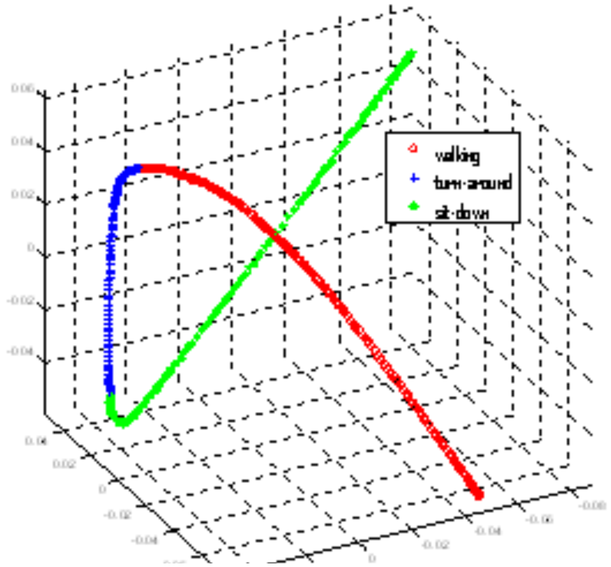


Tong Lin and Hongbin Zha, Riemannian Manifold Learning, IEEE TPAMI 2008

1. MANIFOLD LEARNING

Examples of data manifold: *segmenting a motion sequence*



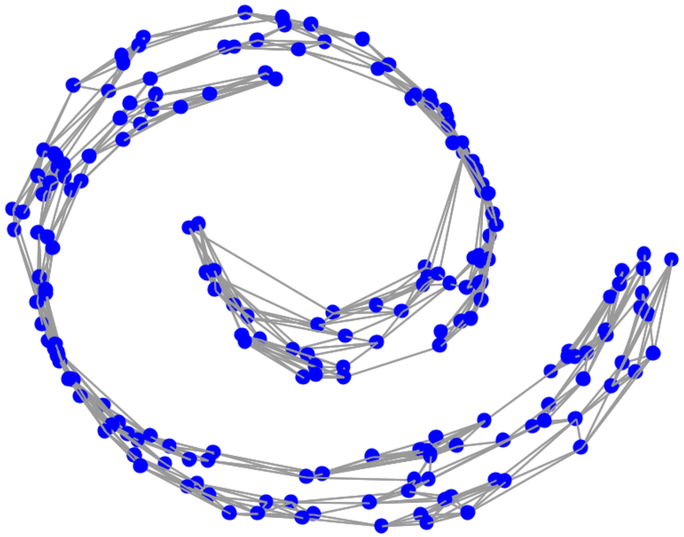


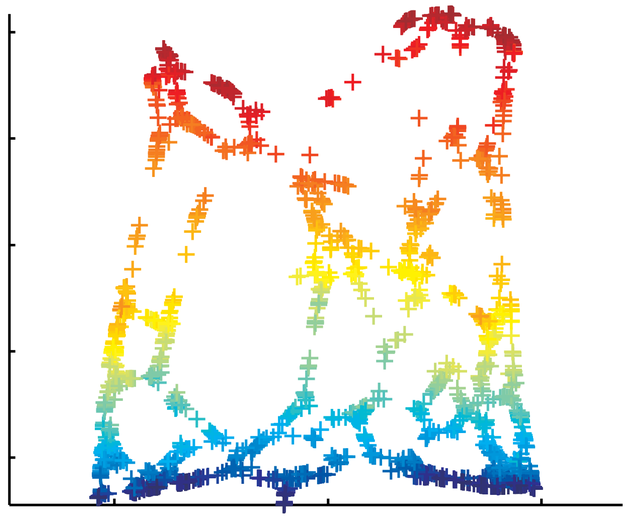
Skeleton joint features embedded in 3D

Dian Gong, Xuemei Zhao, Gérard Medioni, Robust Multiple Manifolds Structure Learning, ICML, 2012

1. MANIFOLD LEARNING

- **Graph-based Manifold Learning (e.g. Laplacian Eigenmaps)**
  - **Use a graph to encode neighborhood information**
  - **Find representation that preserves the local geometry of the data manifold**





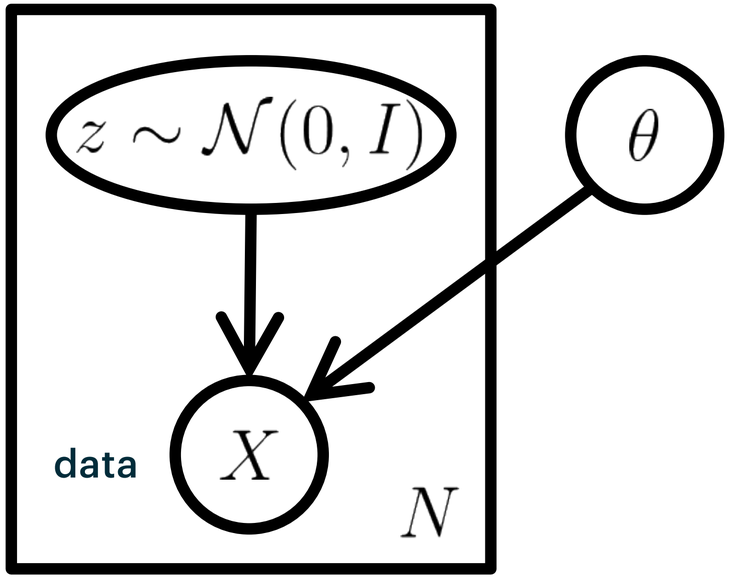
M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *NIPS*, volume 14, pages 585--591, 2001.

2. GENERATIVE MODELS

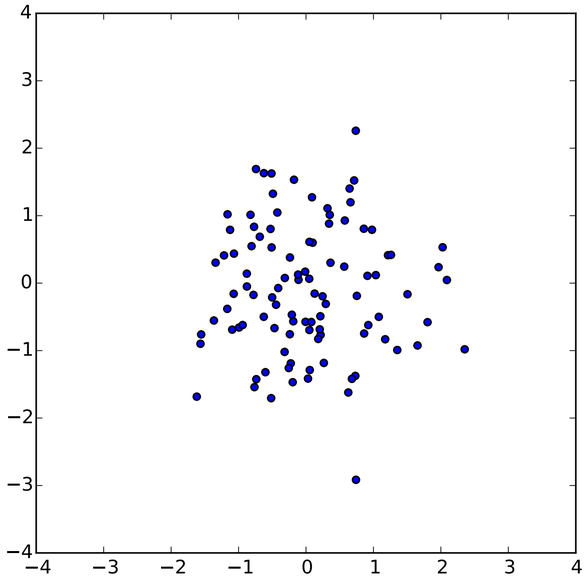
Variational Auto-Encoder (VAE)

A generative approach to non-linear dimension reduction

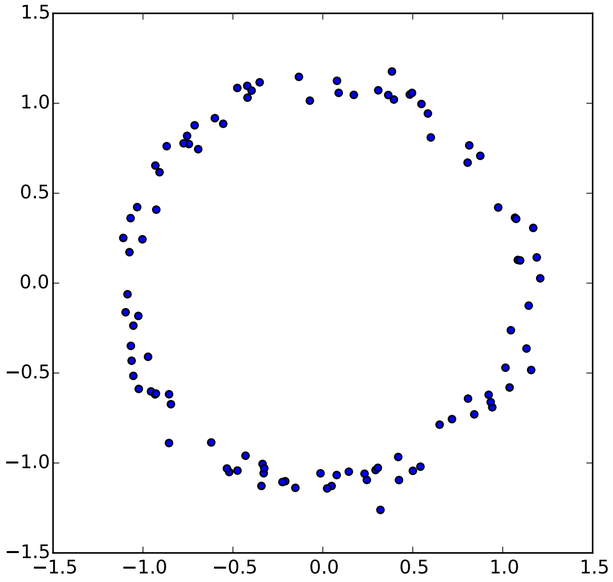
latent variable



$$P(X) = \int P(X|z;\theta)P(z)dz.$$



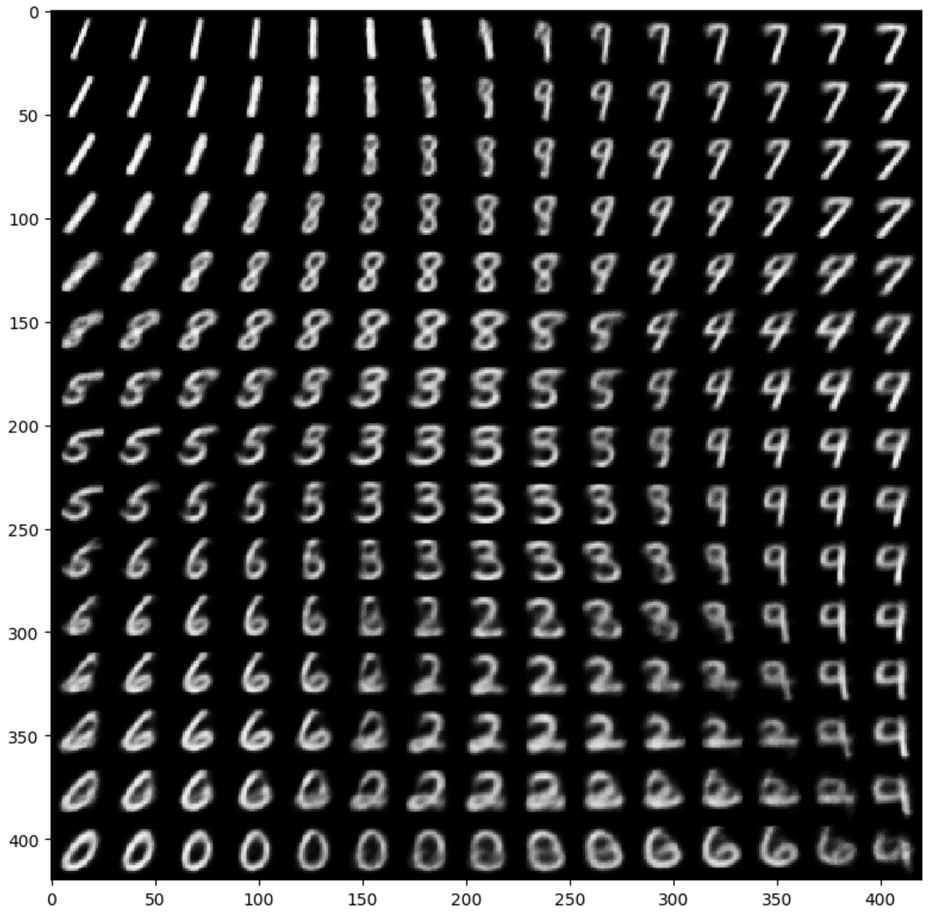
z



$x = z/10 + z/\|z\|$

2. GENERATIVE MODELS

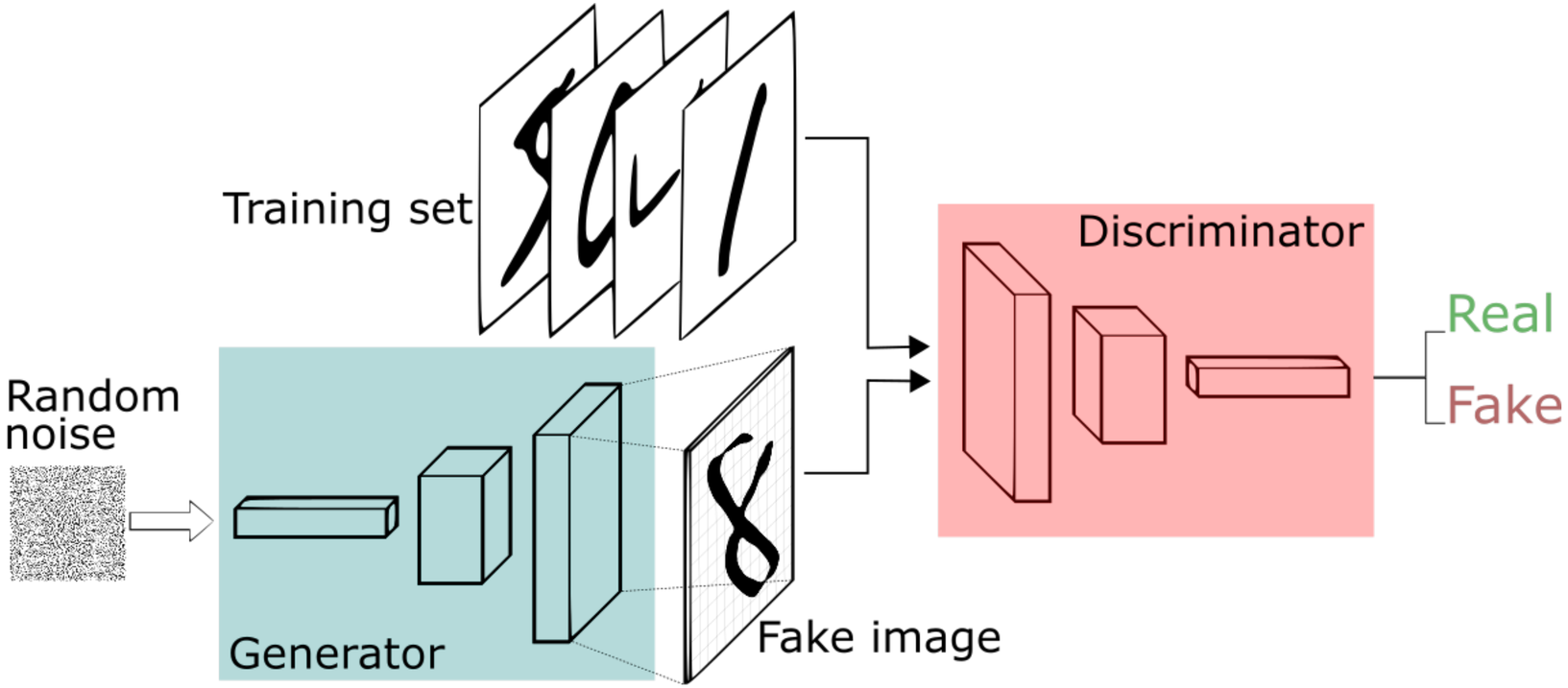
- Variational Auto-Encoder (VAE)
  - Has weak assumptions and small approximation error
  - Can be trained efficiently with backpropagation



MNIST Digits Generated by VAE visualized in 2D

2. GENERATIVE MODELS

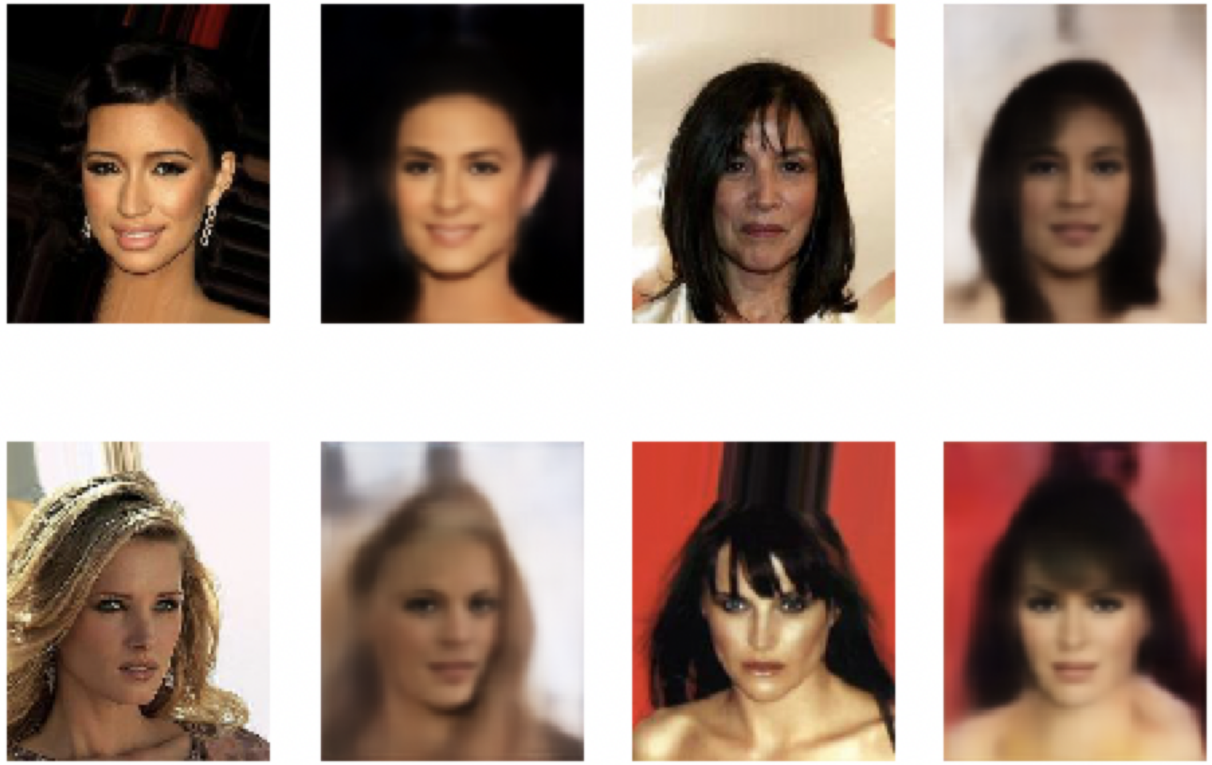
- **Generative models are difficult to approximate**
- **GAN: Generative modeling + Discriminative adversary**



2. GENERATIVE MODELS

GAN can generate more realistic images

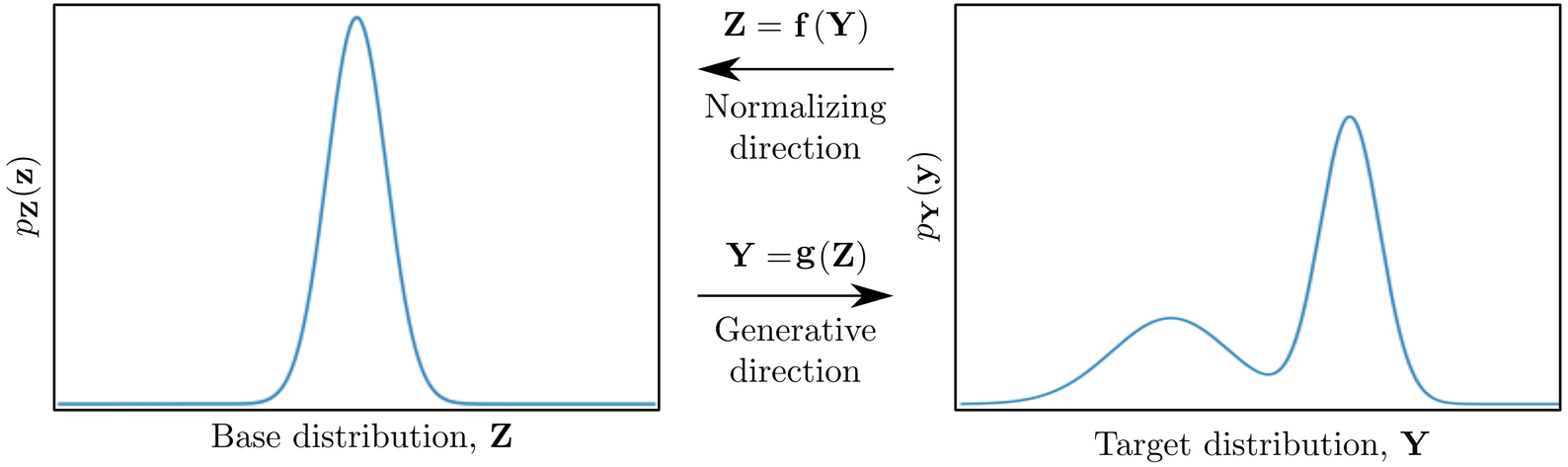
GAN (left) vs VAE (right)



GAN does not provide tractable estimation of density $P(x)$

2. GENERATIVE MODELS

- **Normalizing Flows： learning complex data distribution by transforming a simple density through an invertible map**
  - **allow both sampling and density evaluation.**
  - **Good for representing correlations in output structures**

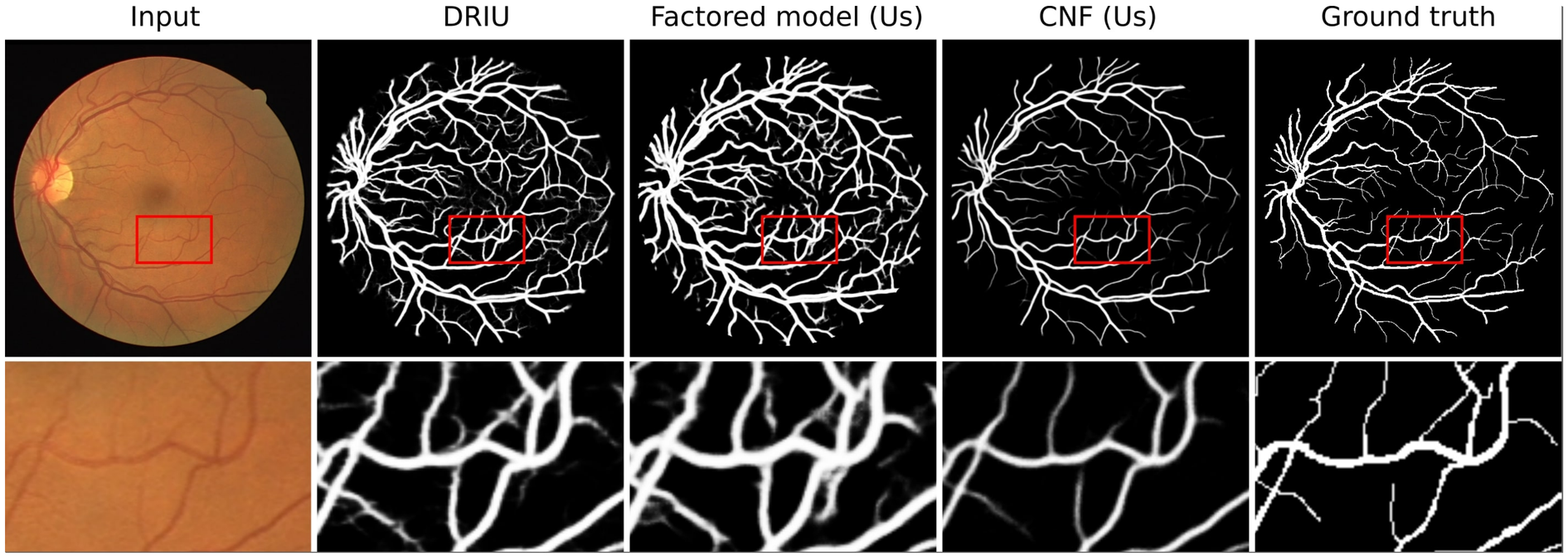


GAN (left) vs VAE (right)

Danilo Jimenez Rezende, Shakir Mohamed, Variational Inference with Normalizing Flows, 2015

2. GENERATIVE MODELS

Example of Conditional Normalizing Flows:

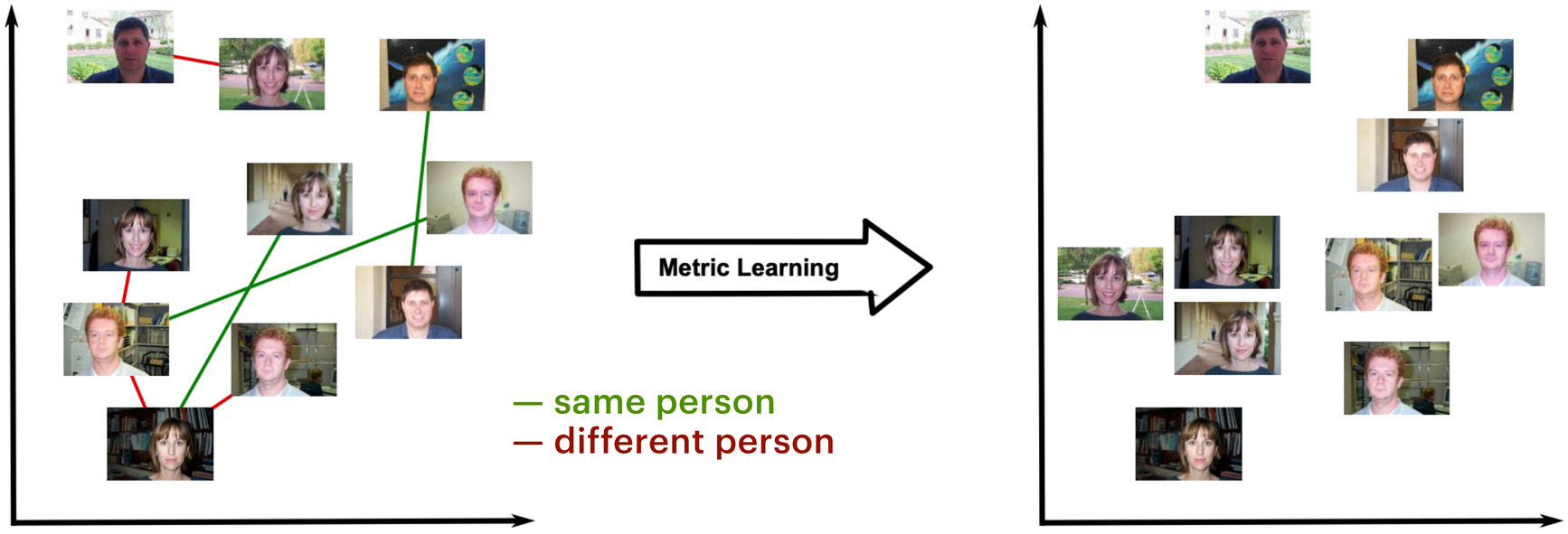


GAN (left) vs VAE (right)

Winkler, Christina, et al. "Learning likelihoods with conditional normalizing flows." arXiv preprint arXiv:1912.00042 (2019).

3. METRIC LEARNING

- **Learn the similarity metric between samples**
- **Useful for few-shot image classification, information retrieval etc**

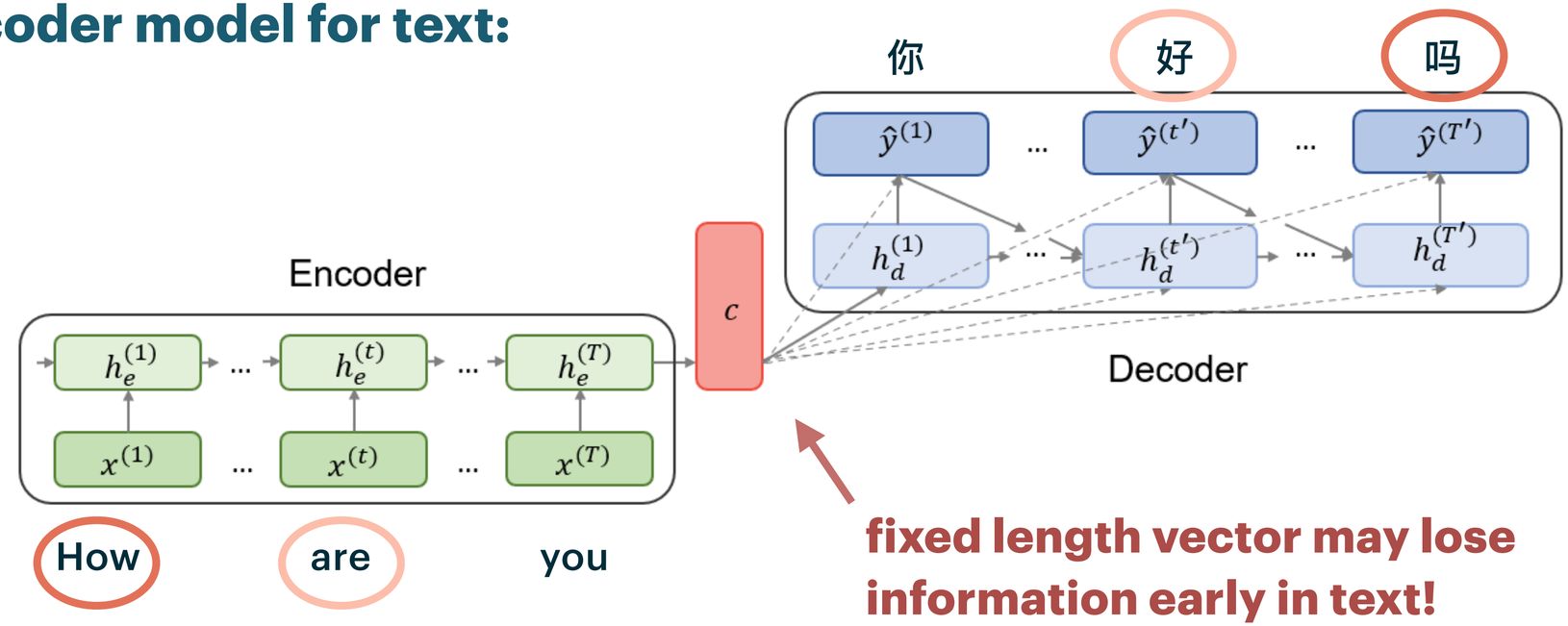


Pairwise constraints showing whether two pictures are the same person

Adapt the metric to the constraints

4. ATTENTION MECHANISM

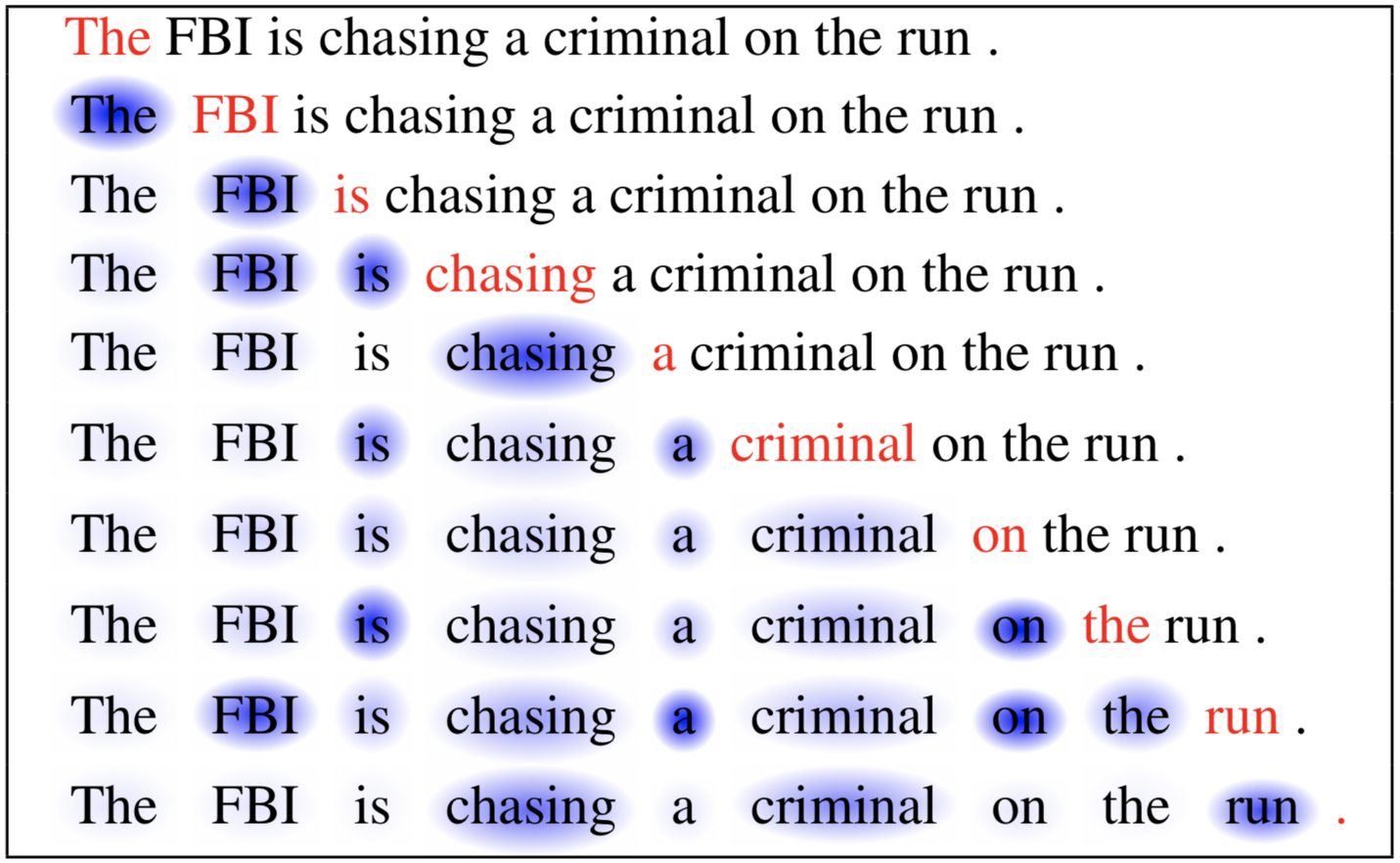
- **Sequence-to-Sequence Problem: map variable-length input sequences to variable-length output sequences (e.g. machine translation)**
- **Encoder-decoder model for text:**



fixed length vector may lose information early in text!

4. ATTENTION MECHANISM

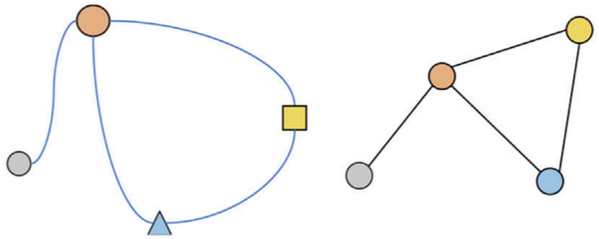
- **Attention enables decoder to "attend" to the whole input sentence.**
- **Self-Attention: A type of attention that relate to different positions of a sentence**



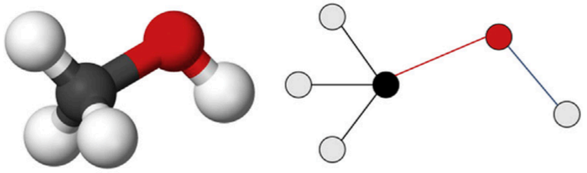
Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

5. GRAPH REPRESENTATION LEARNING

- **Graph neural networks (GNNs): deep learning based methods that operate on graph domain**

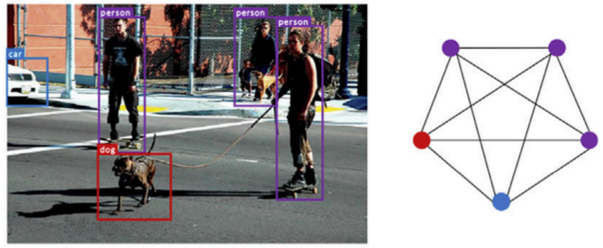


(a) Physics



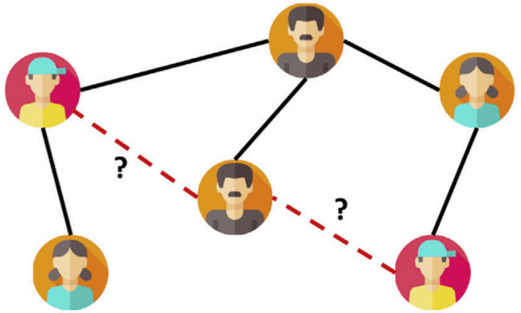
(b) Molecule

Useful for node classification, link prediction, clustering, etc



(c) Image

(d) Text



(e) Social Network

LECTURE OUTLINE

- About this class
- A preview of the topics we will cover
- Break
- Discussion on *Representation Learning: A New Perspective*

DISCUSSION QUESTIONS

- **Discussion on *Representation Learning: A New Perspective***

NEXT CLASS

- **Reading Assignment：** ***Laplacian Eigenmaps for Dimensionality Reduction and Data Representation (Belkin & Niyogi, 2003)*** ***Section 5 is optional***

Recommended group size：3

Feel free to use other references

- **Homework （for non-presenters）**
  - **Reading summary**
  - **Pre-discussion questions**