CS 550: MACHINE LEARNING
Learning

- In our lives, we take actions according to
  - What we observe in our environments
  - What we have previously learned

- Some daily life problems include
  - Face recognition
  - Handwritten character/digit recognition
  - Chess playing
  - Car driving
  - Stock price prediction
Learning

- In order to achieve a task, we should
  - Have relevant information representing the environment
  - Know the possible set of actions
  - Know the process to take an action based on the information
    - This process relies on our past experience
Handwritten letter recognition

- Obtain information representing the environment
  - Letter to be recognized
  - Preferably its adjacent letters

- Know the possible set of actions
  - Number of letters
  - Language

- Take an action, which is affected by whether or not
  - You have seen that letter before
  - You know the alphabet of that language
  - You understand the context of that language
Machine learning

- The goal of machine learning is
  - To design computer systems that automatically achieve tasks, with respect to some performance measures, using past experience
  - To have machines that automatically take actions similar to ours depending on the environment
Machine learning

- Design systems that
  - Automatically take actions (output) similar to ours
  - Depending on the environment (input)
  - Based on their past experience (training samples)
Machine learning

- We reduce the input measuring its certain properties (features), which can be numerical or non-numerical
  - Mileage (e.g., 34187)
  - Condition (e.g., poor, average, excellent)

- The output can be discrete or continuous
  - A, C, Z for letter recognition (classification)
  - Ali, Ayse, Cigdem for face recognition (classification)
  - 25999 TL for car price prediction (regression)
  - 3.7° by which a wheel is turned at each time (regression)
We believe that there is a process underlying training samples (past experience).

- We may not identify this process completely.
- But we can construct a model approximating the process.
  - A function that distinguishes discrete outputs (classification).
  - A functional description of output in terms of inputs (regression).

Machine learning mainly focuses on constructing such models.
Machine learning

- The goodness of the model depends on
  - How well your approximation is
    - No model fits all problems
    - Different models have different assumptions
  - How well training samples represent the distribution in the real-world
    - There may exist noise and exceptions in the samples
    - Some parts may not be covered by the samples

MODEL

input → MODEL → output
How to design a learning system

1. Start
2. Collect data
3. Choose features
4. Train learner
5. Choose model
6. Evaluate learner
7. End
Unsupervised learning

So far, we have talked about **SUPERVISED** learning
- There is a teacher that provides a label (output) for each training sample
- The task is to map an input space to an output space

In **UNSUPERVISED** learning
- There is not explicit teacher that provides outputs
- The task is to find regularities (clusters) in the input space
  - e.g., cluster customers based on their demographic information and past transactions for developing marketing strategies
  - e.g., cluster pixels based on their colors for image compression
An example: Image compression

Image compression to reduce the number of bits to be transferred

RGB color space
→ 24 bits for each pixel

8 (= 2^3) clusters (colors)
→ 3 bits for each pixel

32 (= 2^5) clusters (colors)
→ 5 bits for each pixel
Reinforcement learning

- **REINFORCEMENT** learning is an approach to control learning that accommodates indirect or delayed feedback
  - Training experience is in the form of indirect information consisting of action sequences and final outcome
  - There are no input/output pairs as in the case of supervised learning
  - The environment could be dynamic such that it could be influenced by the selected action
An example: Chess playing

- The system should learn
  - How to choose a sequence of correct actions (moves)
  - In a dynamic environment (chess board)
  - Using past experience (move sequences and final outcomes of various games played)
  - To reach a goal (win chess)