CH. 1: Introduction

- **Traditional AI** makes machines intelligent by manually adding explicit rules to follow, e.g., medical diagnosis systems

  ML makes machines clever by automatically learning rules from a large set of examples. ML is a branch of **AI** since the ability of turning experience into expertise is a manifestation of intelligence.

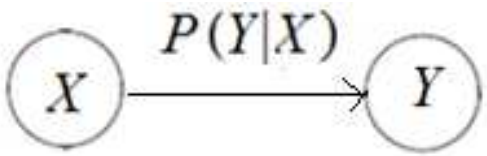
  DL makes machines smart using parallel simple algorithms to extract rules from a large number of examples. DL can be viewed as a branch of ML.

1.1 Why is Machine Learning

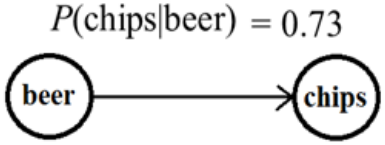
- Widespread use of computers, networks and wireless communication leads to big data
- Data preserves useful information (e.g., relationships, associations, regularities, structures, knowledge, etc.).

Example: Associations – expressed as conditional probabilities $P(Y \mid X)$.

Graphically,



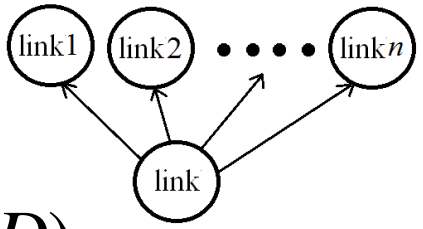
Ex1: Supermarkets – Find association rules between products bought by customers



e.g., $P(\text{chips} \mid \text{beer}) = 0.73$: The customer has the probability of 0.73 of buying chips after buying beer.

Ex2: Web links – Estimate the links a user is likely to click and use this information to download associated pages in advance for faster access,

e.g., $P(\text{link } n \mid \text{link}) = p_n$,

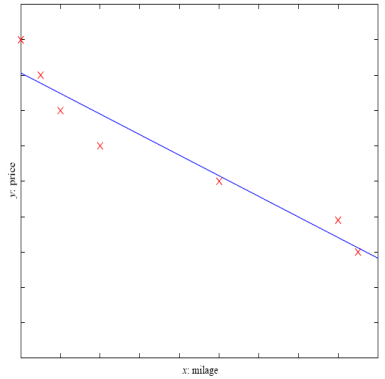


Extensions: $P(Y \mid X, D)$, $P(Y, E \mid X, D)$, ……

Example: Regularity – expressed as functions

$$y = f(x) \text{ or } f(x \mid \boldsymbol{\theta})$$

Ex. Price of car $y = f(x \mid \boldsymbol{\theta})$



x : car attributes, y : price,

$\boldsymbol{\theta}$: parameters

Linear model : $y = a_1 x + a_0,\ \boldsymbol{\theta} = (a_0, a_1)$

Quadratic model :

$$y = a_2 x^2 + a_1 x + a_0,\ \boldsymbol{\theta} = (a_0, a_1, a_2)$$

Gaussian model:

$$y = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right);\ \boldsymbol{\theta} = (\mu,\ \sigma)^T$$

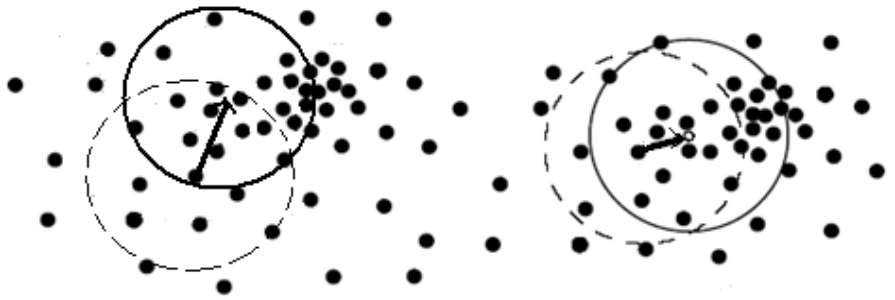
Examples: Structure

Ex1. Density estimation

Data points

$\{x_i\}_{i=1,\cdots\cdots,n}$

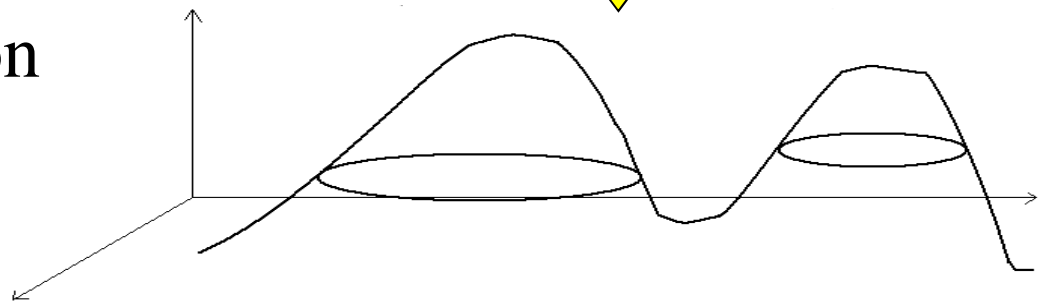
$x_i \in R^d$



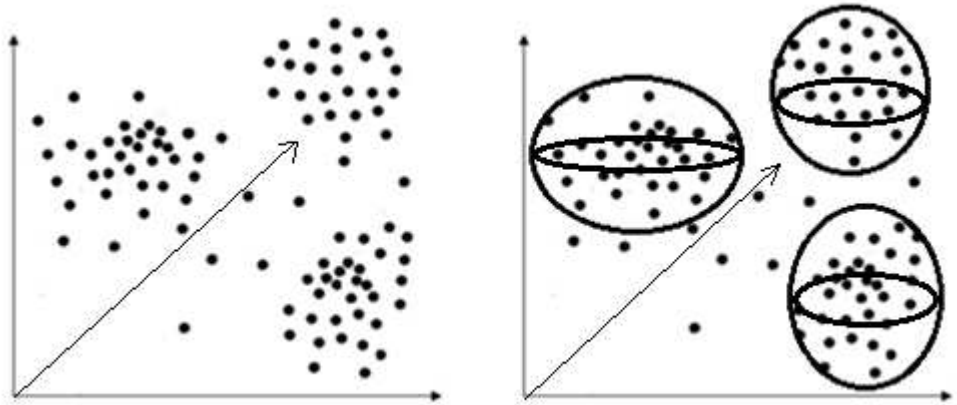
Density Estimation

Density distribution

$p(x)$



Ex2. Clustering– partitioning items into homogeneous subsets



Ex3. Image segmentation

Example: Knowledge

Ex1. Credit scoring – find a classification rule for banks to calculate the risk of a customer given his/her information (e.g., income, collateral, saving, career, etc.) and the amount of credit.

Ex2. Pattern recognition – speech recognition, speaker identification, face recognition, fingerprint, optical character recognition.

- Information extraction may be performed manually through human analysis or automatically through machine learning

- Fundamental ingredients of machine learning:

  Training, validation, test examples

  Hypothesis set – a set of hypothesis models from which the target model is to be determined.

  Inductive bias, Prior knowledge – additional information.

  Learning algorithm – figures out the model that best fits the given training sample.

1.2 Biological Learning

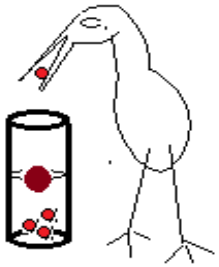
(A) Bacteria, virus

(B) Insects: bee, fly, grasshopper

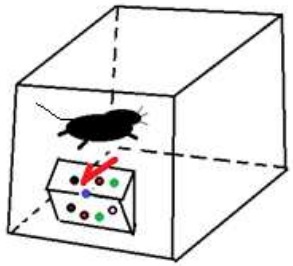
Forage, breed, predator avoidance

(C) Animals:

Bird



Mouse

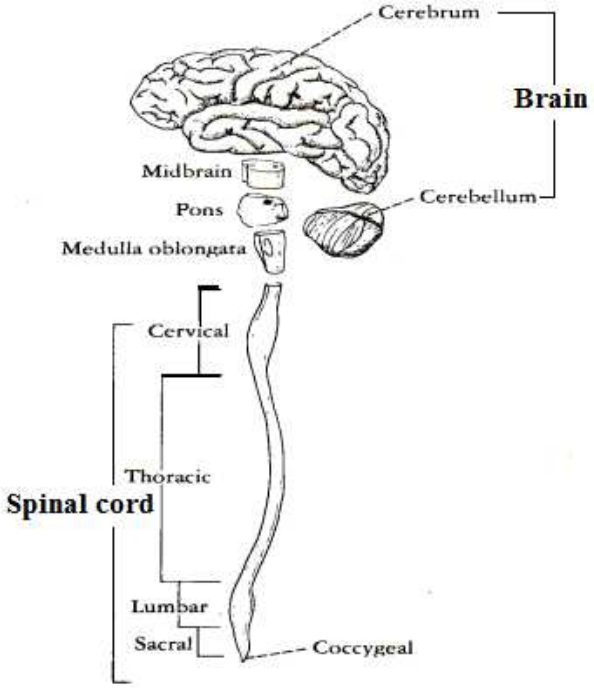


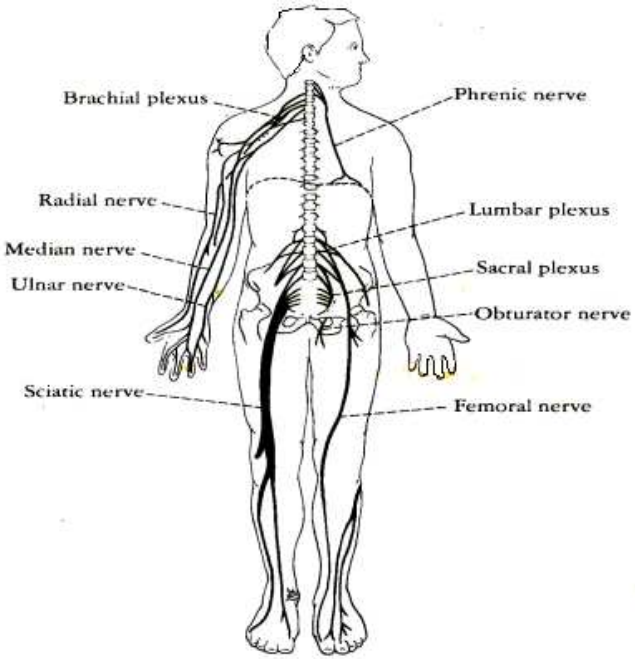
Squirrel

(C) Human:

Learning behavior is primarily due to the nervous system.

Three major components constructing the nervous system: brain, spinal cord, periphery

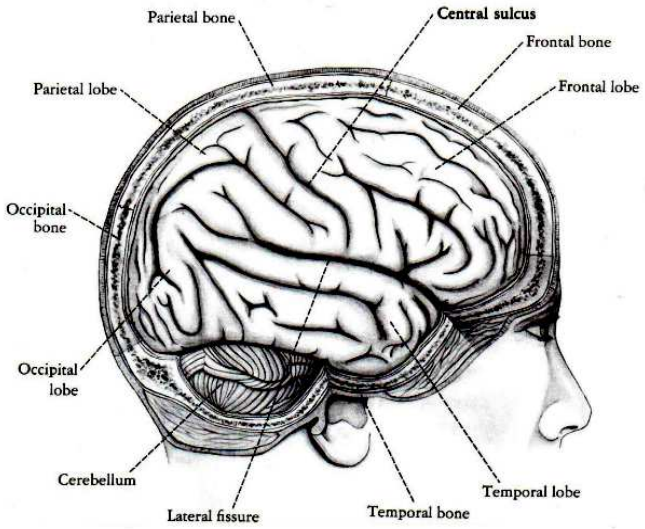




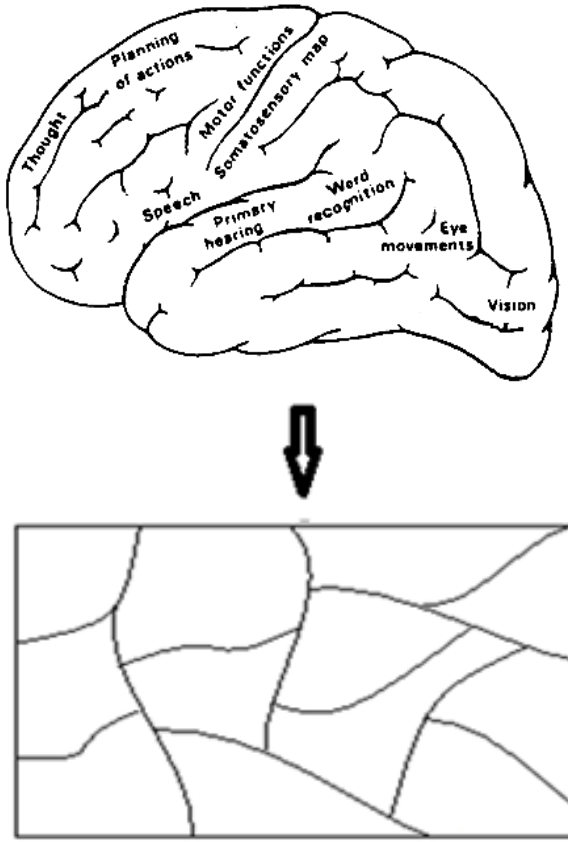
Peripheral nervous system

Learning is rooted at the brain that is the most complex computing device that we have ever known.

Brain



Cerebral cortex



i) Size: $1m^2$

ii) Thick: $2 \sim 4mm$

iii) 6 layers of neurons

- Ordered Feature Map

Macroscopic scale:

e.g. Auditory cortex -- tonotopic maps

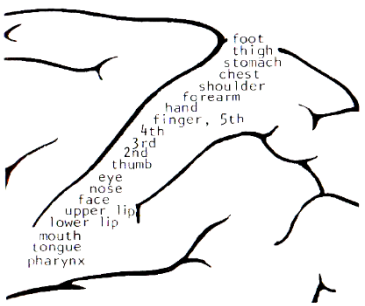
Visual cortex -- retinotectal maps

Somatosensory cortex -- somatic maps

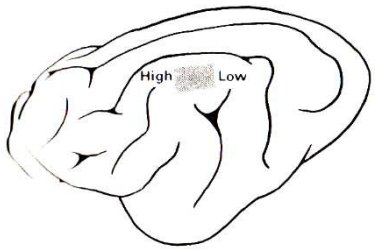
Hippocampal cortex -- geographic maps

Microscopic scale:

e.g. Somatosensory cortex

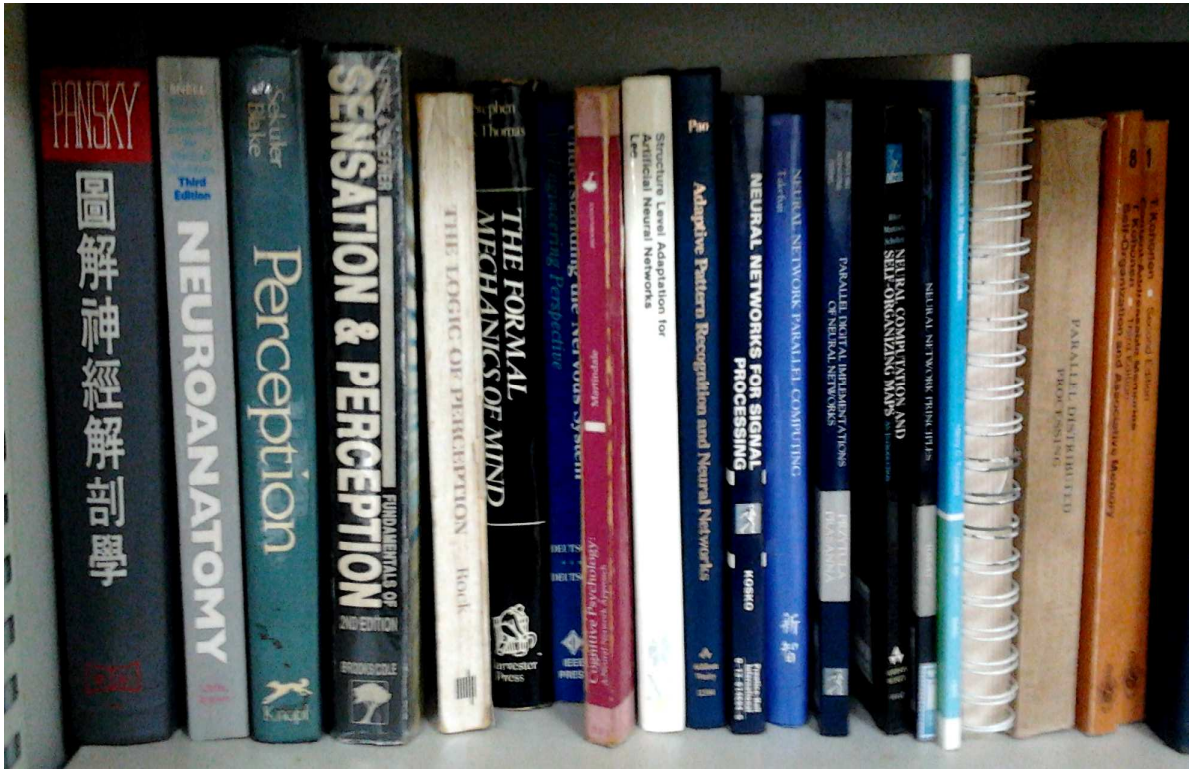


Auditory cortex



Tissue level – inherence, Mental level – learning

Activities: sensation, perception, cognition, recognition, imagination, dream, consciousness.



Sensation -- the process of detecting stimuli in the environment.

Organs: eye, ear, nose, tongue, skin

Perception – the process of converting physical energies into neural patterns.

Cognition – the process by which we generate representations of the world.

Tasks: learn, model, store

Recognition -- re-cognition

Tasks: retrieve, match, decide

Imagination – arrange intentionally a scenario based on the models constructed by cognitive processes

Dream – organize involuntarily a story based on the representations built in cognition

Consciousness -- ……

Unconsciousness -- ……

Various neural models were proposed to mimic the activities of human brain.

- What is the relationship between human learning and machine learning?

1.3 Types of Machine Learning

- Supervised Learning
- Unsupervised Learning
- Semi-supervised Learning
- Self-supervised Learning
- Temporally Supervised Learning
- Reinforcement Learning
- On-line Learning
- Active Learning

Supervised learning -- Learn a mapping from the input to an output using a set of labeled examples

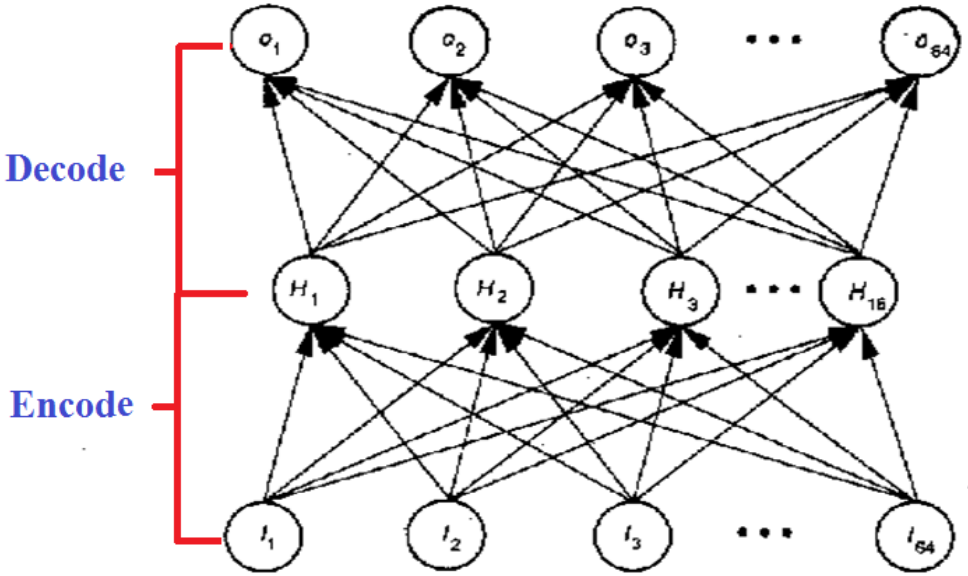
Unsupervised learning -- Find the regularities of data using a set of unlabeled examples, e.g., clustering, density estimation, image segmentation

Semi-supervised Learning -- The training sample consisting of both labeled and unlabeled data

- **Self-supervised Learning**

  -- is a supervised learning without human-annotated labels; they may be generated from the input data or a heuristic algorithm

  Examples: Auto-encoder: the targets are the input



Temporally Supervised Learning

-- Supervision comes from future input

Examples:

i) Predict the stocks of a company

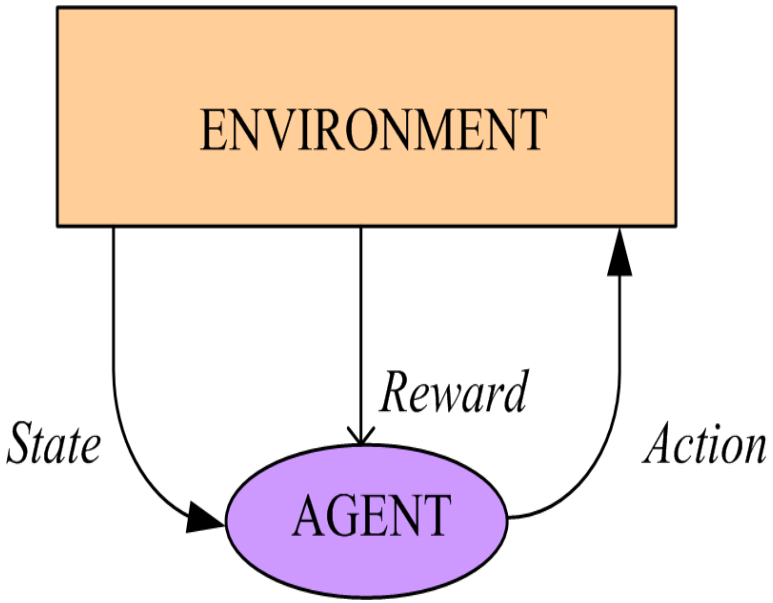
ii) Predict the next frame in a video where the targets are the input

On-line learning -- Training and production phases are intermixed.

Active learning -- The learner adaptively or interactively collects training examples

Reinforcement learning -- Learner through continuously interacting with the environment, from which rewards are accumulated. The goal is to find the optimal policy that maximizes the accumulated reward.

e.g., policy learning,
game playing,
multiple agents,
robot navigation,
credit assignment.



- Fields relate to machine learning:

  i) Artificial intelligence, ii) Artificial neural networks

  iii) Data mining -- knowledge discovery in databases

  iv) Generative models -- explain the observed data through the interaction of hidden factors

  v) Graphical models -- visualize the interaction of hidden factors and data

  vi) Bayesian formalism -- define prior knowledge on the hidden factors and models

  vii) Clustering, classification, pattern recognition