1 Introduction

- **Training Set** : is used to tune the **parameters** of an adaptive model

  What about adaptive model??

  How to design or select an good adaptive model??

- **Learning phase** : The precise form of the function(adaptive model) is determined during the *training phase*

- **General Workflow**

  1. Preprocess:
     - **Normalization** : For example : transforming different size of pics to the same one.
     - **Reduction** : Reduce dimension or quantity, to speed up comuputation. How to reduce dimension/quantity while preserving the information in those data is not easy.
  2. Training
  3. Test : We can't train our model on all data that we have. We have to separate some of them to verify our model.
  4. Output result

Classification:

- **supervised learning** : Training data comprise examples of the input vectors along with their corresponding target vectors
  - **classification**:decrete
  - **regression**:continuous

- **unsupervised learning** : Training data consists a set of input vector $\mathbf{x}$ without any corresponding target value
  - **cluster** : to discover group of similar examples
  - **density estimation** : to determine the distribution of data within the input space.

- **visualization** : to project the data from a high-dimensional space down to two/three dimensions for the purpose of *visualization*. Data visualization is a hot field now. Is this can be used to pre-process data set,as mention in *Reduction*?

- **reinforcement learning**: finding suitable actions to take in a given situation in order to maximize a reward.
  Make a balance of exploration and exploitation

  - **exploration** : explore the unknow space.
  - **exploitation** : make use of the actions that are known to yield a high reward.

  Do you remember PSO ???

  $$V(t+1) = w*V(t)+C_1*R_1*(P(t)-X(t))+C_2*R_2*(G(t)-X(t)) \quad (1)$$

  $$X(t+1) = X(t) + V(t+1) \quad (2)$$

  It also need a balance between exploitation and exploration.But it has a global attraction when particle explore the unknown space. It is lucky...HA!

- **deep learning** : a new branch of machine learning. wikipedia

2 Mathematica

Covariance

$$cov(X, Y) = E((X - \mu)(Y - \nu)) \quad (3)$$

Why?Why covariance is define like this?We can see what it is defined in :

> Covariance provides a measure of the strength of the correlation between two or more sets of random variates.

1. It is a measure of correlation. **correlation** means two or more factors are equal. They play the role. So they should be the **same form** in covariance equation....

2. It is product of the difference between item and expectation. Product is a good choice. It can refect the relationship perfectly.

The variance of a random variable X is its second central moment, the expected value of the squared deviation from the mean

$$\sigma^2(X) = E[(X - \mu)(X - \mu)] \tag{4}$$

The variance can also be thought of as the covariance of a random variable with itself:

$$cov(X, Y) = E((X - \mu)(Y - \mu)) \tag{5}$$

Normal Distribution A normal distribution in a variate X with mean μ and variance σ^2 is a statistic distribution with probability density function

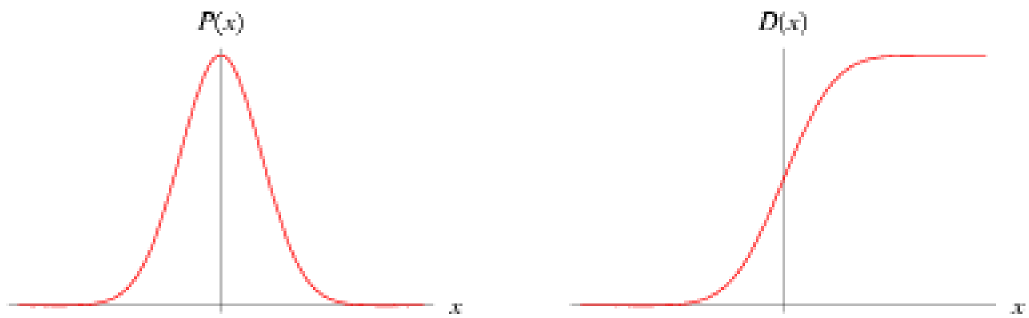

Figure 1: Normal Distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \tag{6}$$

Why **Normal Distribution** is useful?According to Central Limit Theorm, the **mean** of any set of variates with any distribution having a finite mean and variance tends to the normal distribution.

$$X_{norm} \equiv \frac{\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu_i}{\sqrt{\sum_{i=1}^{N} \sigma^2}} \tag{7}$$

Then ???Why this is useful? What X_{norm} means? What is the **normal form**?