③ 判斷 function 好 / 壞 : Loss function

④ 設模型找 Best function
  - 1. Gradient ascent
    - 1. Regularization
    - 2. Generative model : Max. Likelihood
    - 3. Cross entropy
  - 2. Gradient descent
    - 1. Feature scaling
    - 2. Backpropagation
    - 3. Backpropagation Through Time (BPTT 算法)
    - 4. Momentum
  - 3. Max. Likelihood 之均值和標準差
    - 1. New activation function
      - 1. train RBM
      - 2. Maxout (ReLU...)
    - 2. Adaptive Learning rate - RMSProp
      - ↳ RMSProp + Momentum = Adam
    - 3. Dropout : ensemble 精神及用法

☑ 補充 : 1. Deep Learning 之 Train 技巧
  - 1. training set 不好
    - 1. Early Stopping
    - 2. Regularization
  - 2. testing set 不好
    - (此才叫 Overfitting)
    - 3. Dropout

(應用) ⇒ word embedding
  - 1. Count based
  - 2. Prediction based

2. Deep (同瘦, 多層) v.s. Shallow (肥矮, 1層)
  ⇒以證出 Deep > Shallow, 即為相同之模組化概念!
  (同層模組化概念)

3. 關於 Dimension Reduction
  - Linear
    - 1. PCA
    - 2. Autoencoder { 1. Encoder
                     { 2. Decoder
    - 3. NMF
  - nonlinear { 1. LLE, 2. Laplacian Eigenmaps, 3. t-SNE