GENERALISED RADIAL BASIS FUNCTION NETWORKS

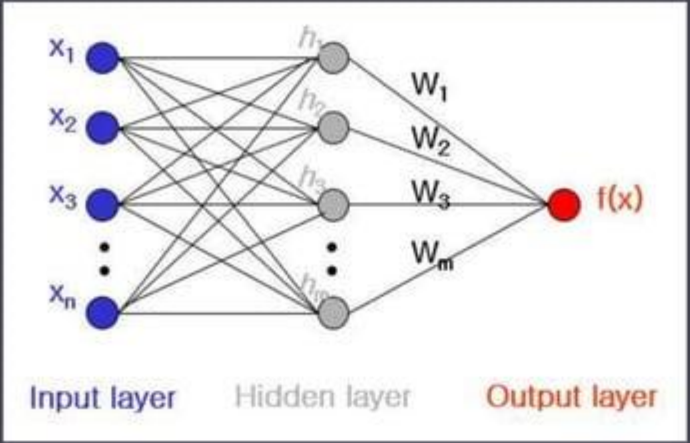
What is Radial basis Function Network???

- Artificial neural network that uses **non-linear radial basis functions** as activation functions.
- Gives **linear output** using combination of **radial basis functions of the inputs and neuron parameters**.
- Useful for function approximation, time series prediction, classification and system control.

Features

- Two-layer feed-forward networks.
- Hidden nodes implement a set of radial basis functions (e.g. Gaussian functions).
- Output nodes implement linear summation functions as in an MLP.
- Training/learning is very fast.
- Networks are very good at interpolation.

Architecture



Radial Basis Function

$$f(x) = \sum_{j=1}^{m} w_j h_j(x)$$

$$h_j(x) = \exp(-(x-c_j)^2 / r_j^2)$$

Where c_j is center of a region,

r_j is width of the receptive field

Commonly Used Radial Basis Functions

Gaussian Functions:

$$\phi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

width parameter $\sigma > 0$

Multi-Quadric Functions:

$$\phi(r) = \left(r^2 + \sigma^2\right)^{1/2}$$

parameter $\sigma > 0$

Generalized Multi-Quadric Functions:

$$\phi(r) = \left(r^2 + \sigma^2\right)^{\beta}$$

parameters $\sigma > 0,\ 1 > \beta > 0$

Inverse Multi-Quadric Functions:

$$\phi(r) = \left(r^2 + \sigma^2\right)^{-1/2} \qquad \text{parameter } \sigma > 0$$

Generalized Inverse Multi-Quadric Functions:

$$\phi(r) = \left(r^2 + \sigma^2\right)^{-\alpha} \qquad \text{parameters } \sigma > 0, \alpha > 0$$

Thin Plate Spline Function:

$$\phi(r) = r^2 \ln(r)$$

Cubic Function:

$$\phi(r) = r^3$$

Linear Function:

$$\phi(r) = r$$

Cover's Theorem

- A complex pattern-classification in a high-dimensional space nonlinearly instead of linear separation in a low-dimensional space.

  e.g. considering X-OR gate.

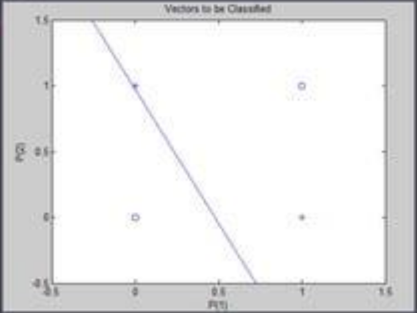


Fig1. Separation using MLP.

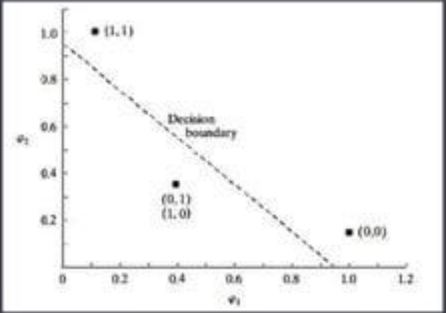


Fig2. Separation using RBF.

Interpolation Problem

- The ***exact interpolation*** of a set of N data points in a multi-dimensional space requires every one of the D dimensional input vectors $\mathbf{x}^p = \{x_i^p : i = 1,\ldots, D\}$ to be mapped onto the corresponding target output t^p.
- The goal is to find a function $f(\mathbf{x})$ such that,

$$f(\mathbf{x}^p) = t^p \qquad ; p = 1,\ldots, N$$

- For strict interpolation, the interpolating surface (i.e. function f) is constrained to pass through all the training data points.

Training/Learning of RBF Network

- Center and spread learning (or determination)
  - Fixed centers selected at random
  - Self-organized selection of centers
  - Supervised selection of centers
- Output layer Weights Learning

Fixed Centers Selected At Random

- Fixed RBFs of the hidden units.
- Locations of the centers may be chosen randomly from the training data set.
- Can use different values of centers and widths for each radial basis function.
- Only output layer weight is need to be learned.
- Obtain the value of the output layer weight by pseudo-inverse method.
- Require a large training set for a satisfactory level of performance

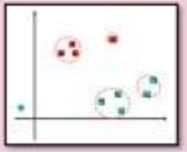
Self-organized Selection of Centers

- Hybrid learning
  - self-organized learning to estimate the centers of RBFs in hidden layer
  - supervised learning to estimate the linear weights of the output layer
- Self-organized learning of centers by means of clustering.
  - k-means clustering
    1. Initialization
    2. Sampling
    3. Similarity matching
    4. Updating
    5. Continuation
- Supervised learning of output weights by LMS algorithm.

Supervised Selection of Centers

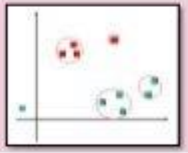
- All free parameters of the network are changed by supervised learning process.
- Error-correction learning using LMS algorithm

Comparison of RBF and MLP

RBF	MLP
Local receptive field. (Only inputs near a receptive field produce an activation.)	Global hyper plane. (All inputs cause an output.)
Larger number of hidden neurons.	Smaller number of hidden neurons.
Longer computation time.	Shorter computation time.
Shorter learning time.	Longer learning time.
Curve fitting in RBF:- 	Curve fitting in MLP:- 


Thank
You

Comparison of RBF and MLP

RBF	MLP
Local receptive field. (Only inputs near a receptive field produce an activation.)	Global hyper plane. (All inputs cause an output.)
Larger number of hidden neurons.	Smaller number of hidden neurons.
Longer computation time.	Shorter computation time.
Shorter learning time.	Longer learning time.
Curve fitting in RBF:- 	Curve fitting in MLP:- 


Thank
You