Machine Learning
10. week

- Mapping Concept
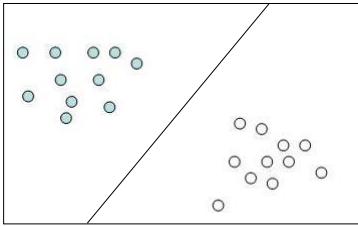- Radial Basis Functions (RBF)
- RBF Networks

Umut ORHAN, PhD.

1

1

Mapping

It is probably the best scenario for the classification of two dataset is to separate them linearly. As you see in the below figure, two different dataset are classified linearly.

But this scenario is not always possible in real life datasets.

Umut ORHAN, PhD.

2

2

Mapping

All of the proposed machine learning methods have been developed using linear classification.

So they can not find solution for the non linear dataset.

Instead of solving datasets in their own space, data is moved to a new space where they can be separated linearly.
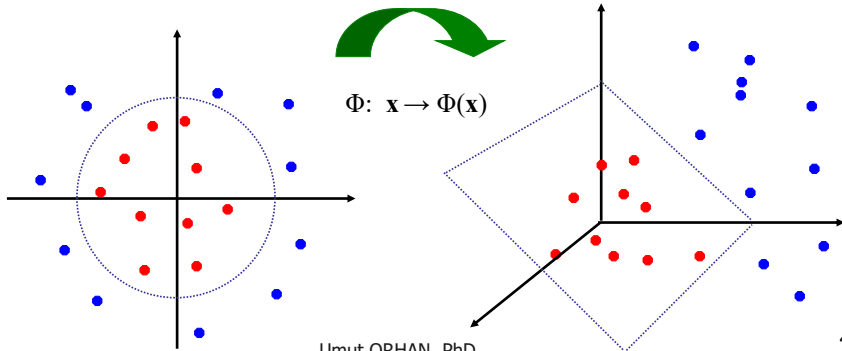
This process is called Mapping.

Umut ORHAN, PhD.

3

3

Mapping

Generally a Φ function is used to convert x into a linear seperatable status.

$$\Phi: \mathbf{x} \rightarrow \Phi(\mathbf{x})$$

Umut ORHAN, PhD.

4

4

Mapping

Mapping Function
- Polynomial
- Radial basis
- Sigmoid

- Implicit mapping to a higher dimensional space where linear separation is possible.

Complicated separation boundary

Simple separation boundary: Hyperplane

Umut ORHAN, PhD.

5

5

Mapping with RBF

Φ Function RBF (Radial Basis Function) is given equation below. Where (c_j) is the center points which represents data, (x_i) is the data points.

Φ Function is the exponantial affect of the distance between (c_j) and (x_i)

$$\Phi(x) = \exp\left(-\frac{\left\|c_j - x_i\right\|^2}{2r^2}\right)$$

Umut ORHAN, PhD.

6

6

RBF Networks

Neurons in hidden layer are calculated as in below equation, where x_i is the data point from input with no weight, c_j is the prototype hidden in the neuron. Simply it is a Euclidian distance.

$$d = \left\| c_j - x_i \right\|$$

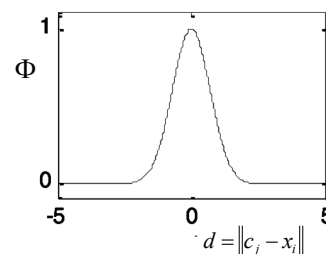After that this distance is evaluated in Radial Basis Function.

Umut ORHAN, PhD.

7

7

RBF Networks

Widely used radial basis function equation and graphical relation of distance-function output is shown below.
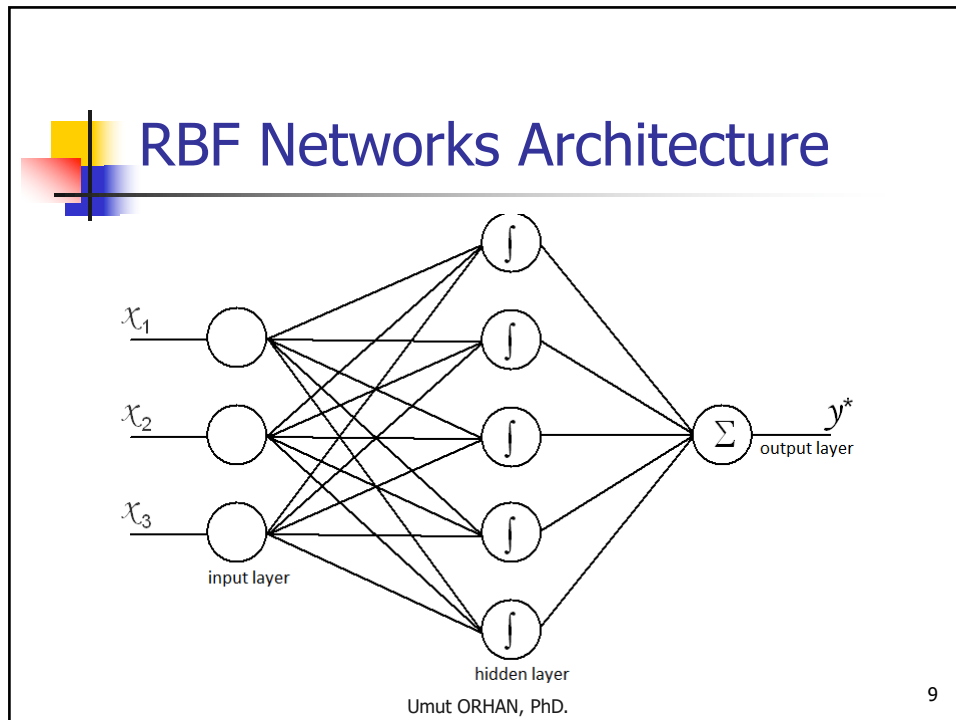
$$\Phi(x) = \exp\left(-\frac{\left\| c_j - x_i \right\|^2}{2r^2} \right)$$

$$d = \left\| c_j - x_i \right\|$$

r value represents the diameter of prototype.

Umut ORHAN, PhD.

8

8

RBF Networks Architecture

x_1

x_2

x_3

input layer

hidden layer

$\int$ $\int$ $\int$ $\int$ $\int$

Σ → y^*

output layer

Umut ORHAN, PhD.

9

9

Estimation in RBF Networks

y represents target class, and it can be calculated with following equation where *x* represents input.

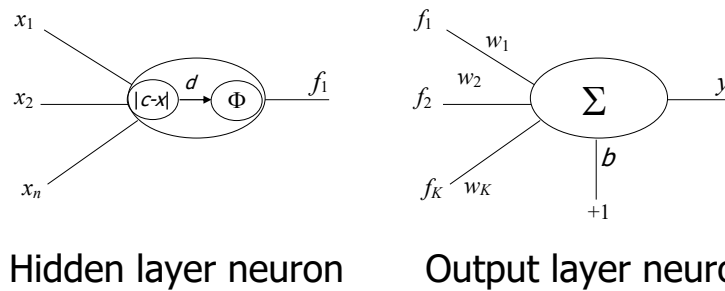$$y_j = \sum_{1 \le i \le K} w_i \, \Phi(x_j)$$

K value shows number of neuron in hidden layer.

Umut ORHAN, PhD.

10

10

Structure of RBF Networks

In below figures you can see neurons of hidden layer and output layer.

Hidden layer neuron   Output layer neuron

Umut ORHAN, PhD.

11

11

Features of RBF Networks

- RBF networks are feed forward networks with supervised learning.
- Generally a unique hidden layer is used in which each neuron contents a RBF function.
- Although training of RBF networks looks similar with back-propogation, they are trained faster than MLP.
- With advances of radial basis fuctions, they are less affected from unstable input problems.

Umut ORHAN, PhD.

12

12

Training in RBF Networks

There are four parameters in RBF networks which is unknown and need to be learned:

- Number of neurons in hidden layer
- Coordinate of each prototype neuron
- Prototype diameter of each neuron
- Output Weights

Umut ORHAN, PhD.

13

13

Training in RBF Networks

In most of the proposed methods, it is outlined that number of neurons in hidden layer variable needs to be researched by trying.

Since output weights are components of a linear equation, solution is easy and it depends on neuron outputs.

So, most important point in training depends on to know position and diameter values of prototypes.

Umut ORHAN, PhD.

14

14

Training in RBF Networks

Prototype position information in neurons can be found by unsupervised learning methods (like K-Means)

- Training can be faster by running supervised and unsupervised trainings paralelly.
- Accuracy of unsupervised learning classifiers are always lower than fully supervised learning methods.

Umut ORHAN, PhD.

15

15

Training in RBF Networks

Another proposal is to use the randomly selected samples as a prototype and not update position information. In this case, diameter (variants) data should be known and selected samples should represent every region in data space very well.

Umut ORHAN, PhD.

16

16

Training in RBF Networks

In order to find Prototype diameter (width of each RBF unit, also known as spread) K-nearest neighbour algorithm is used.

Root-mean squared distance between current cluster center and K nearest neighbours is calculated. And this is the value chosen for the unit width(r).

So if current cluster center is cj, the r value is;

$$rj = \sqrt{\sum_{i=1}^{k} \frac{(cj - ci)^2}{k}}$$

Umut ORHAN, PhD.

17

17

Training in RBF Networks

When we study structure of RBF networks, we can find some similarities with MLP network model. By means of linear structured similarity in output layer, it might be benefited from back propagation algorithm during training of RBF networks.

$$E = \frac{1}{2} \sum_{j} e_j^{\,2} \qquad \Delta w_i = -\eta \frac{\partial E}{\partial w_i}$$

Umut ORHAN, PhD.

18

18

Training in RBF Networks

Output weights are updated with back propagation training and update quantities of function exits in hidden layer neurons can be predicted. This can be used indirectly in update of prototype diameter values and protoype coordinates.
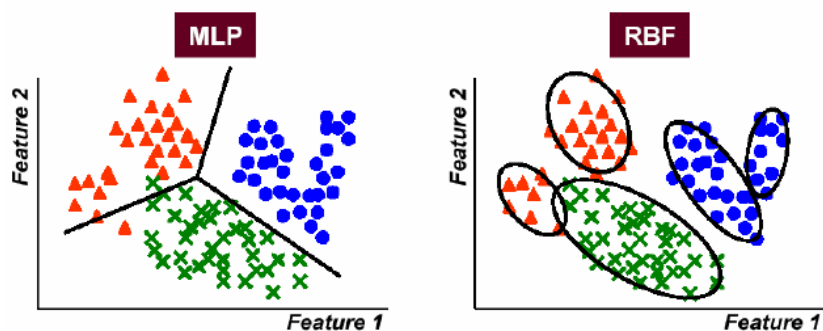
$$\Delta c_i = -\eta \frac{\partial E}{\partial c_i} \qquad \Delta r_i = -\eta \frac{\partial E}{\partial r_i}$$

Umut ORHAN, PhD.

19

19

Classification with RBF Networks

Umut ORHAN, PhD.

20

20

Presentation Task

Please compare success of the results using linear regression for following mapping functions by using two dimensional non-linear artificial dataset.

- Sigmoid
- Hiperbolic Tangent
- RBF
- 3th degree polynomial

Umut ORHAN, PhD.

21

21