Radial Basis functions: (RBF)

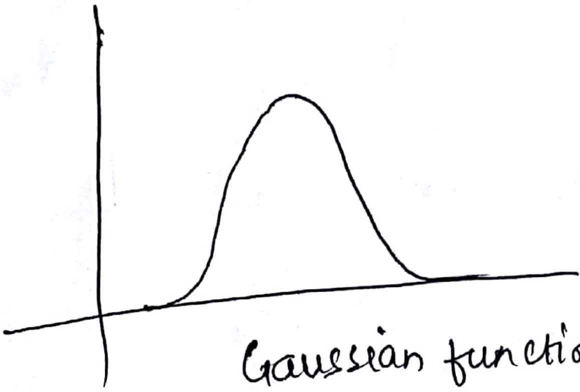
RBF are a special class of functions for nonlinear models. The basis of RBF is based on cover's theorem. As per cover's theorem, a nonlinear problem can be linearly seperable when the problem is elevated to higher dimensional space. This feature monotonically increase or decrease with distance from a central point.

- The center, the distance, the shape of the radial function are the parameters of the model.
- A typical radial function which is invariably used is Gaussian function which in the case of a scalar input is:

$$h(x) = \exp\left(-\frac{(x-c)^2}{r^2}\right)$$

where c = center, r = radius

- A gaussian RBF monotonically decreases with the distance from the center or in the other words the function has a single peak at the center and the edges tapering monotonically on both ends.



Gaussian function.

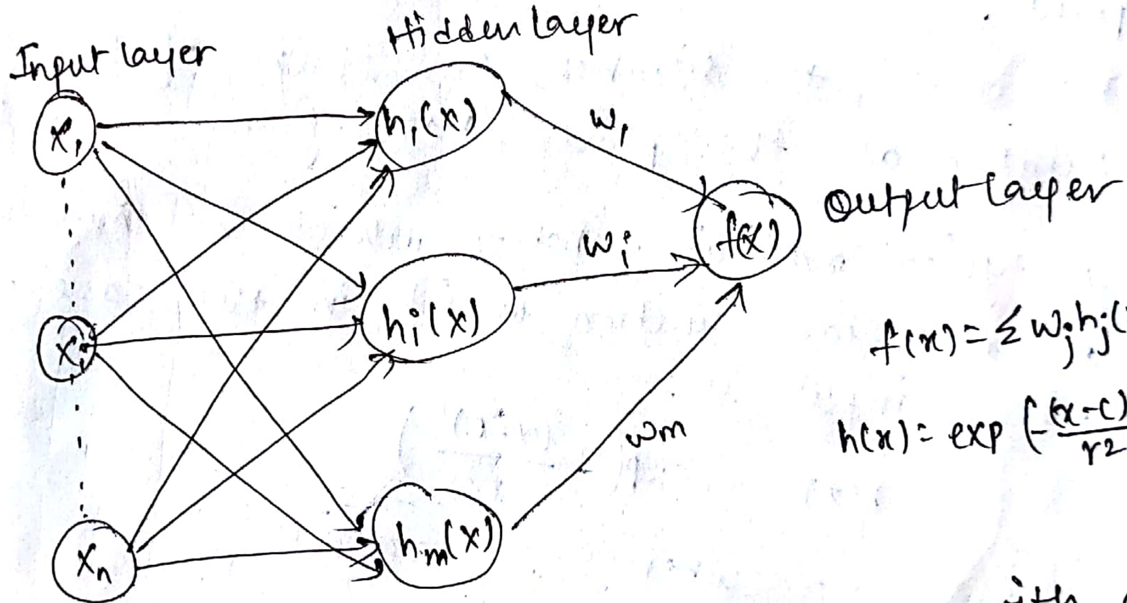
Radial Basis Function Networks :

→ This sort of functions can be employed in any sort of model, which can be either linear or nonlinear.

→ It can also be employed with a single layer or multilayered network.

→ The architecture of RBF has three layers – an Input layer, a hidden layer, and a output layer.

→ Basic Architecture of RBF is shown as-



$$f(x) = \sum w_j h_j(x)$$

$$h(x) = \exp\left(-\frac{(x-c)^2}{r^2}\right)$$

The focus here is on a single layer network with a single hidden layer where nonlinear mappings of higher dimensionality takes place.

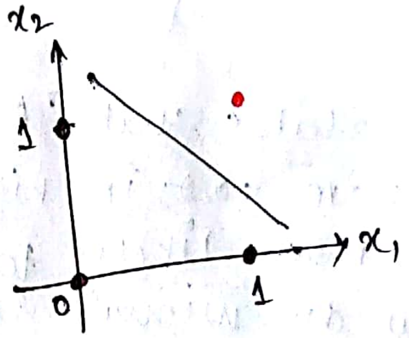
Scanned by CamScanner

Single layer RBF Network :-

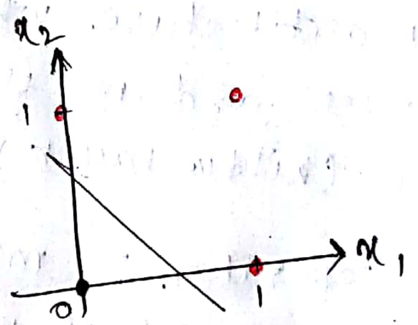
In a single perceptron, we can establish linear seperability, as in the case of AND and OR functions. for AND and OR functions, the outputs are either 1 or 0 and outputs can be linearly seperable. But this is not the case with EXOR function.

→ This can be understood dramatically -

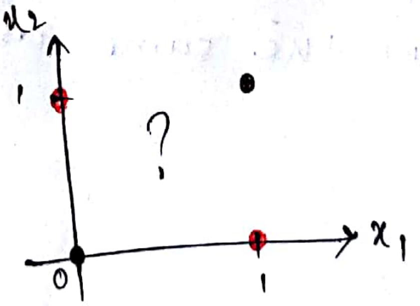
x_1	x_2	x_1 AND x_2
0	0	0 (Black dot)
0	1	0
1	0	0
1	1	1 (Red)



x_1	x_2	x_1 OR x_2
0	0	0 (Black)
0	1	1 Red
1	0	1
1	1	1



x_1	x_2	x_1 EXOR x_2
0	0	0
0	1	1
1	0	1
1	1	0



EXOR gate is not linearly seperable in contrast to AND gate and OR gate.

→ for seperating the input patterns, we need atleast one hidden layer.

→ The RBF network (RBFN) transforms the input single into another form and this is fed into the network to get linear seperability. RBFN is structurally same as perceptron.

Scanned by CamScanner

→ RBFN is composed of input, hidden and output layers.

→ If RBFN is having exactly one hidden layer, then this hidden layer is known as feature vector.

→ we apply nonlinear transfer function to the feature vector before we solve the classification problem.

→ when we increase the dimension of the feature vector, the linear seperability of feature vector increases.

Cover's Theorem →

cover's theorem states that "A complex pattern-classification problem cast in high-dimensional space non-linearly is more likely to be linearly seperable than in a low dimensional space."

- consider an RBFN architecture where n is the no. of input features/values and m is the no. of transformed vector dimensions (hidden layer width)

- for nonlinearity seperation: m≥n. Each node in the hidden layer performs a set of nonlinear an RBF.

- The output remains the same as for the classification problems.

Eg: Construct an RBF pattern classifier such that

1) (0,0) and (1,1) are mapped to class 0, class C_1

2) (1,0) and (0,1) ———————— 1, class C_2

Sol^n

① centers are selected at random. $\mu_1 = (0,0)$, $\mu_2 = (1,1)$

② Activation function of hidden neuron is computed as follows:

$$\phi_i(\|x\|) = \exp\left[-\frac{m_1}{d_{max}^2}\|x-\mu_i\|^2\right]$$

where m_1 = no of centers

d_{max} = max. dist between two centers.

$m_1 = 2$

$d_{max} = \sqrt{(0-1)^2 + (0-1)^2} = \sqrt{2} = 1.414$

$\phi_1(x) = \exp(-\|x-\mu_1\|^2)$

$\phi_2(x) = \exp(-\|x-\mu_2\|^2)$

x_1	x_2	$\phi_1(x)$	$\phi_2(x)$
0	0	1.000	0.1353
0	1	0.3679	0.3679
1	0	0.3679	0.3679
1	1	0.1353	1.000

Plotting them,