Contents

1 Model 1: RBFNN model with updates of weights and centres

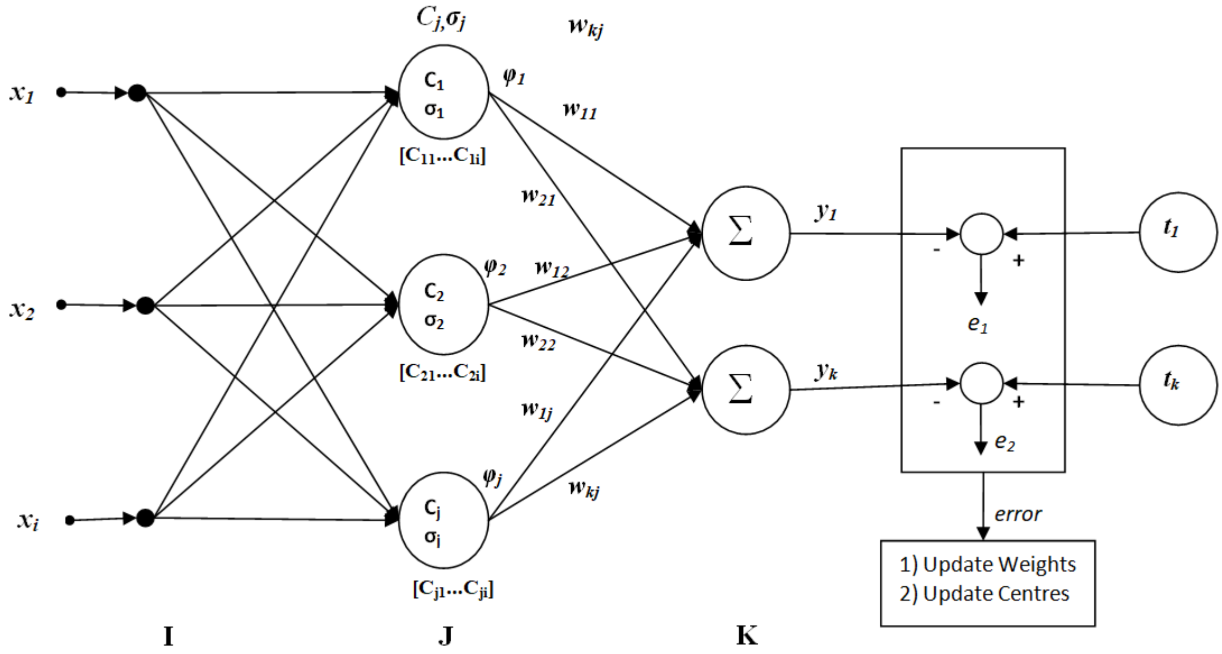

Figure 1: RBFNN model with updates of weights and centres

1.1 Key Equations

- $y_k = \sum_j w_{kj}\phi_j$
- $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$
- $z_j = ||X - C_j|| = \sqrt{\sum_j (x_i - c_{ij})^2}$
- Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$

1.2 Gradient Descent Learning

- $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$
  where η_w = learning rate for update of weights
- $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$
  where η_c = learning rate for update of centres

1.3 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \tag{1}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum_k (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k) \end{aligned} \tag{2}$$

From key equations, we know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial w_{kj}}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial w_{kj}} &= \frac{\partial(\sum_j (w_{kj}\phi_j)}{\partial w_{kj}} \\ &= \phi_j \end{aligned} \tag{3}$$

Using equations (2) & (3) we can rewrite equation (1) as:

$$\begin{aligned} \frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k)\phi_j \end{aligned} \tag{4}$$

1.4 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \tag{5}$$

From equation (2), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$
From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial \phi_j} &= \frac{\partial(\sum_j w_{kj}\phi_j}{\partial \phi_j} \\ &= w_{kj} \end{aligned} \tag{6}$$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial z_j}$ will be:

$$\begin{aligned} \frac{\partial \phi_j}{\partial z_j} &= \frac{\partial e^{\frac{-z_j^2}{2\sigma_j^2}}}{\partial z_j} \\ &= -\frac{z_j \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^2} \\ &= -\frac{z_j \times \phi_j}{\sigma_j^2} \end{aligned} \tag{7}$$

From key equations, We know that $z_j = ||X - C_j|| = \sqrt{\sum_j (x_i - c_{ij})^2}$, so $\frac{\partial z_j}{\partial c_{ji}}$ will be:

$$\begin{aligned}\frac{\partial z_j}{\partial c_{ji}} &= \frac{\partial(\sqrt{\sum_i (x_i - c_{ji})^2})}{\partial c_{ji}} \\ &= -\frac{2 \times (x_i - c_{ji})}{2 \times \sqrt{\sum_i (x_i - c_{ji})^2}} \\ &= -\frac{(x_i - c_{ji})}{z_j}\end{aligned} \tag{8}$$

Using equations (2), (6), (7) & (8) we can rewrite equation (5) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{ji}} &= [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \\ &= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{-(z_j \times \phi_j)}{\sigma_j^2} \times \frac{-(x_i - c_{ji})}{z_j} \\ &= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})\end{aligned} \tag{9}$$

1.5 Final equations of $\frac{\partial E}{\partial w_{kj}}$ $\frac{\partial E}{\partial c_{ji}}$

- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
- $\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} = [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})$

2 Model 2: RBFNN model with updates of weights, centres and standard deviations

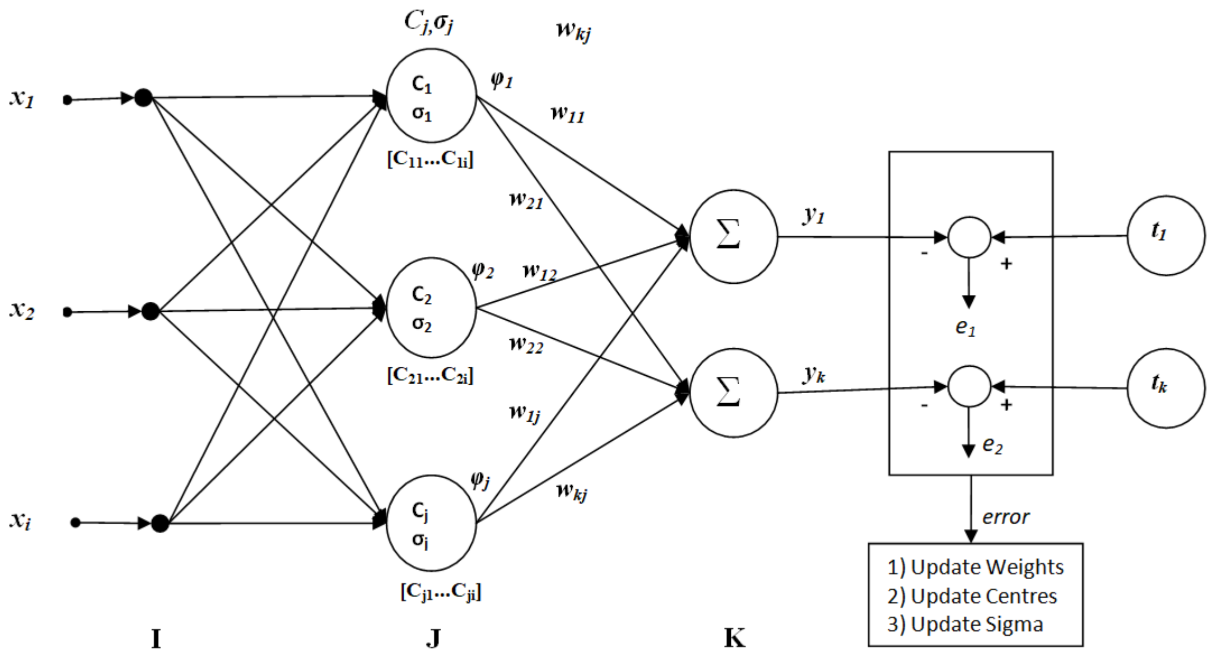

Figure 2: RBFNN model with updates of weights, centres and standard deviations

2.1 Key Equations

- $y_k = \sum_j w_{kj}\phi_j$
- $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$
- $z_j = ||X - C_j|| = \sqrt{\sum_j (x_i - c_{ij})^2}$
- Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$

2.2 Gradient Descent Learning

- $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$
  where η_w = learning rate for update of weights
- $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$
  where η_c = learning rate for update of centres
- $\sigma_{ji}(t+1) = \sigma_{ji}(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_{ji}}$
  where η_σ = learning rate for update of sigmas

2.3 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \tag{10}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum_k (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{11}$$

From key equations, we know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial w_{kj}}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial(\sum_j (w_{kj}\phi_j)}{\partial w_{kj}} \\ &= \phi_j\end{aligned} \tag{12}$$

Using equations (2) & (3) we can rewrite equation (1) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k)\phi_j\end{aligned} \tag{13}$$

2.4 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \tag{14}$$

From equation (2), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$
From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\begin{aligned}\frac{\partial y_k}{\partial \phi_j} &= \frac{\partial(\sum_j w_{kj}\phi_j}{\partial \phi_j} \\ &= w_{kj}\end{aligned} \tag{15}$$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial z_j}$ will be:

$$\begin{aligned}\frac{\partial \phi_j}{\partial z_j} &= \frac{\partial e^{\frac{-z_j^2}{2\sigma_j^2}}}{\partial z_j} \\ &= -\frac{z_j \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^2} \\ &= -\frac{z_j \times \phi_j}{\sigma_j^2}\end{aligned} \tag{16}$$

From key equations, We know that $z_j = ||X - C_j|| = \sqrt{\sum_j (x_i - c_{ij})^2}$, so $\frac{\partial z_j}{\partial c_{ji}}$ will be:

$$\begin{aligned}\frac{\partial z_j}{\partial c_{ji}} &= \frac{\partial(\sqrt{\sum_i (x_i - c_{ji})^2})}{\partial c_{ji}} \\ &= -\frac{2 \times (x_i - c_{ji}}{2 \times \sqrt{\sum_i (x_i - c_{ji})^2}} \\ &= -\frac{(x_i - c_{ji})}{z_j}\end{aligned} \tag{17}$$

Using equations (15), (16), (17) & (18) we can rewrite equation (14) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{ji}} &= [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \\ &= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{-(z_j \times \phi_j)}{\sigma_j^2} \times \frac{-(x_i - c_{ji})}{z_j} \\ &= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})\end{aligned} \tag{18}$$

2.5 Derivation of $\frac{\partial E}{\partial \sigma_j}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial \sigma_j}$ as:

$$\frac{\partial E}{\partial \sigma_j} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}]\frac{\partial \phi_j}{\partial \sigma_j} \tag{19}$$

From equation (11), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$.

From equation (15), we know that $\frac{\partial E}{\partial \phi_j} = w_{kj}$. From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial \sigma_j}$ will be:

$$\begin{aligned}\frac{\partial \phi_j}{\partial \sigma_j} &= \frac{\partial e^{\frac{-z_j^2}{2\sigma_j^2}}}{\partial \sigma_j} \\ &= \frac{z_j^2 \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^3} \\ &= \frac{z_j^2 \times \phi_j}{\sigma_j^3}\end{aligned} \tag{20}$$

Using equations (11), (15) & (20) we can rewrite equation (19) as:

$$\begin{aligned}\frac{\partial E}{\partial \sigma_j} &= [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}]\frac{\partial \phi_j}{\partial \sigma_j} \\ &= [\sum_k -(t_k - yk)w_{kj}]\frac{z_j^2 \times e^{\frac{-x_j^2}{2\sigma_j^2}}}{\sigma_j^3}\end{aligned} \tag{21}$$

2.6 Final equations of $\frac{\partial E}{\partial w_{kj}}$ $\frac{\partial E}{\partial c_{ji}}$

- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
- $\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j}] . \frac{\partial \phi_j}{\partial z_j} . \frac{\partial z_j}{\partial c_{ji}} = [\sum_k -(t_k - y_k) w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})$
- $\frac{\partial E}{\partial \sigma_j} = [\sum_k \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j}] \frac{\partial \phi_j}{\partial \sigma_j} = [\sum_k -(t_k - yk) w_{kj}] \frac{z_j^2 \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^3}$

3 Model 3: RBFNN model with updates of weights in both layers and centres

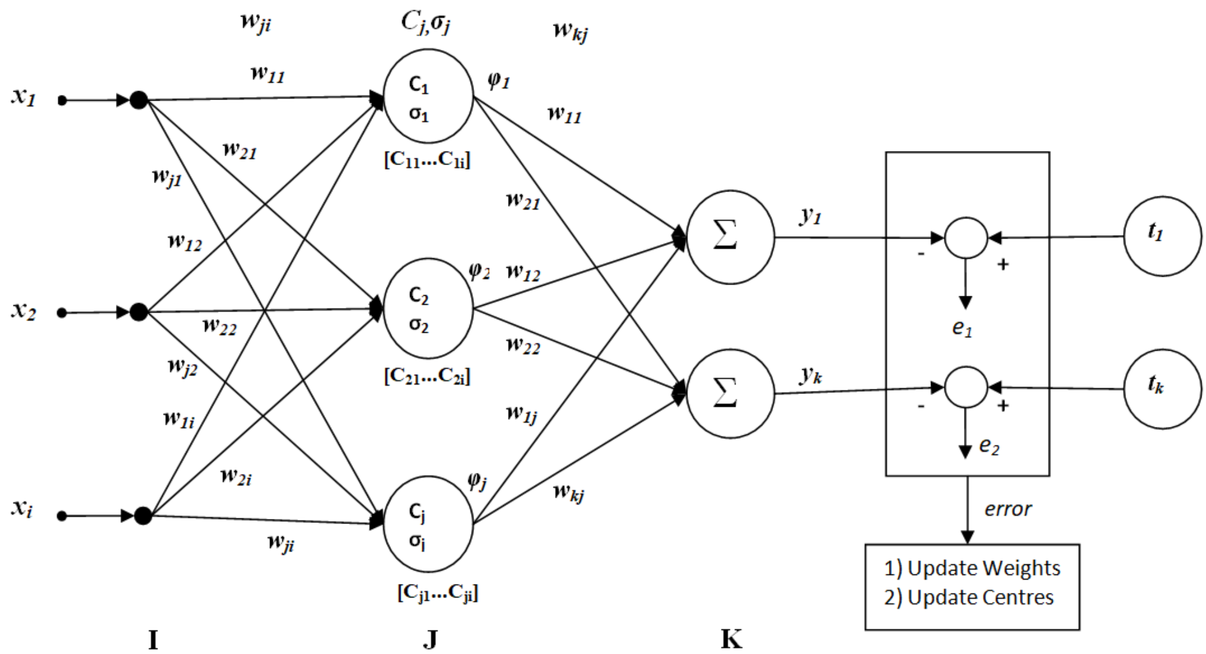

Figure 3: RBFNN model with updates of weights in both layers and centres

3.1 Key Equations

- $a_I = \sum\limits_{i=1}^{I} x_i w_{ji}$
- $z_j = (||a_I - c_j||)$
  Let $||a_I - c_j|| = z =$ Euclidean distance
- $\phi_z = e^{\frac{-z_j^2}{2\sigma^2}}$
- $y_k = \sum\limits_{j=1}^{J} \phi_j w_{kj}$
- Cost function: $E = \frac{1}{2}\sum(t_k - y_k)^2$

3.2 Gradient Descent Learning

- Learning of centres:
$$c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$$
  where η_c = learning rate for update of centres
- Learning of weights:
  For w_{kj},
$$w_{kj}(t+1) = w_{kj}(t) - \eta_{w_1} \frac{\partial E}{\partial w_{kj}}$$

For w_{ji},

$$w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2}\frac{\partial E}{\partial w_{ji}}$$

where η_{w_1} and η_{w_2} = learning rate for update of weights

3.3 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \tag{22}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_- y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_- y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{23}$$

From key equations, we know that $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y_k}{\partial \phi_j} = w_{jk} \tag{24}$$

The equation $\frac{\partial \phi_j}{\partial c_j}$ can be written as

$$\begin{aligned}\frac{\partial \phi_j}{\partial c_j} &= \frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_j} \\ &= \frac{-zj}{\sigma^2}\phi_j.\frac{-(a_j - c_j)}{z_j}\end{aligned} \tag{25}$$

Using equations (23), (24) & (25) we can rewrite equation (22) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{ji}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \\ &= \frac{-(t_k - y_k)w_{kj}\phi_j(a_j - c_j)}{\sigma^2}\end{aligned} \tag{26}$$

3.4 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k).\phi_j\end{aligned} \tag{27}$$

3.5 Derivation of $\frac{\partial E}{\partial w_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{ji}}$ as:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} \tag{28}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{29}$$

From key equations, we know that $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y_k}{\partial \phi_j} = w_{jk} \tag{30}$$

The equation $\frac{\partial \phi_j}{\partial a_I}$ can be written as

$$\begin{aligned}\frac{\partial \phi_j}{\partial a_I} &= \frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial a_I} \\ &= \frac{-zj}{\sigma^2}\phi_j.\frac{-(a_I - c_{ji})}{z_j}\end{aligned} \tag{31}$$

From key equations, we know that $a_I = \sum_{i=1}^{I} x_i w_{ji}$, so $\frac{\partial a_I}{\partial w_{ji}}$ will be:

$$\frac{\partial a_I}{\partial w_{ji}} = x_i \tag{32}$$

Using equations (29), (30), (31) & (32) we can rewrite equation (28) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{ji}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} \\ &= \frac{-(t_k - y_k)w_{kj}\phi_j(a_I - c_{ji})x_i}{\sigma^2}\end{aligned} \tag{33}$$

3.6 Final equations of $\frac{\partial E}{\partial c_{ji}}$, $\frac{\partial E}{\partial w_{kj}}$, $\frac{\partial E}{\partial w_{ji}}$

- $\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k)w_{kj}\phi_j(a_j - c_j)}{\sigma^2}$
- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}}$
  $= -(t_k - y_k).\phi_j$
- $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} = \frac{-(t_k - y_k)w_{kj}\phi_j(a_I - c_{ji})x_i}{\sigma^2}$

4 Model 4: RBFNN model with updates of weights in both layers, centres and standard deviations

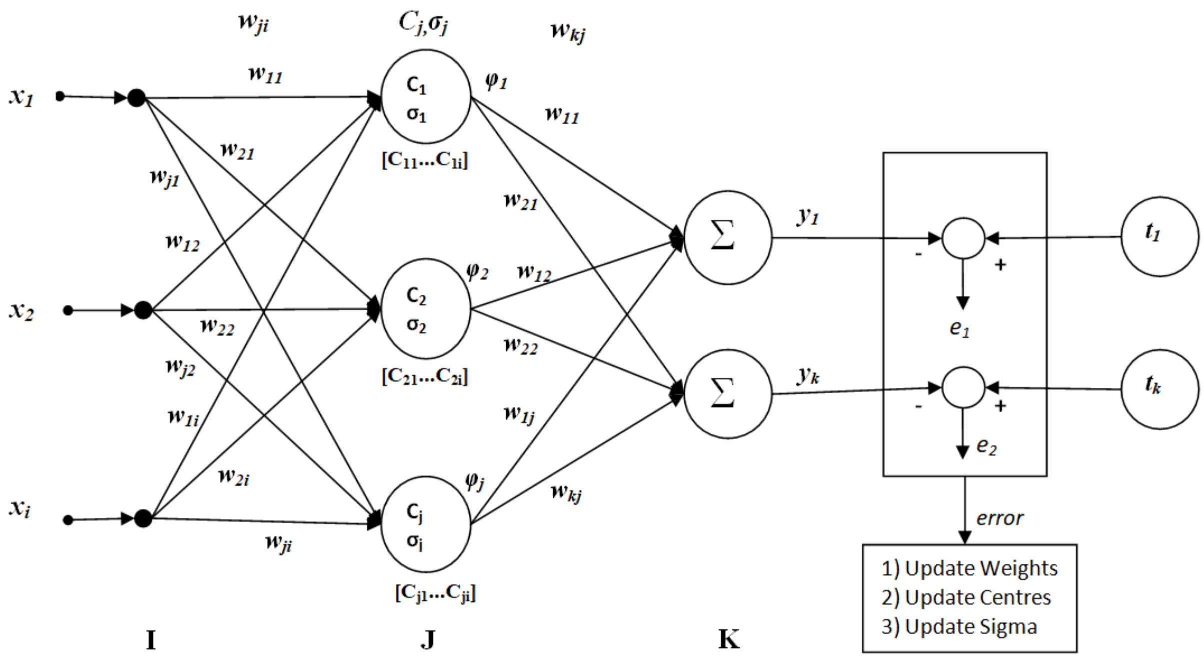

Figure 4: RBFNN model with updates of weights in both layers, centres and standard deviations

4.1 Key Equations

- $a_I = \sum_{i=1}^{I} x_i w_{ji}$
- $z_j = (||a_I - c_j||)$
  Let $||a_I - c_j|| = z =$ Euclidean distance
- $\phi(z) = e^{\frac{-z_j^2}{2\sigma^2}}$
- $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$
- Cost function: $E = \frac{1}{2} \sum (t_k - y_k)^2$

4.2 Gradient Descent Learning

- Learning of centres:
  $$c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$$
  where $\eta_c =$ learning rate for update of centres
- Learning of weights:
  For w_{kj},
  $$w_{kj}(t+1) = w_{kj}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{kj}}$$

For w_{ji},

$$w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{ji}}$$

where η_{w_1} and η_{w_2} = learning rate for update of weights

- Learning of sigma:

$$\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$$

where η_σ = learning rate for update of sigmas

4.3 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \tag{34}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_k - y)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k) \end{aligned} \tag{35}$$

From key equations, we know that $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y_k}{\partial \phi_j} = w_{jk} \tag{36}$$

The equation $\frac{\partial \phi_j}{\partial c_j}$ can be written as

$$\begin{aligned} \frac{\partial \phi_j}{\partial c_j} &= \frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_j} \\ &= \frac{-zj}{\sigma^2}\phi_j.\frac{-(a_j - c_j)}{z_j} \end{aligned} \tag{37}$$

Using equations (35), (36) & (37) we can rewrite equation (34) as:

$$\begin{aligned} \frac{\partial E}{\partial c_{ji}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \\ &= \frac{-(t_k - y_k)w_{kj}\phi_j(a_j - c_j)}{\sigma^2} \end{aligned} \tag{38}$$

4.4 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\begin{aligned} \frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k).\phi_j \end{aligned} \tag{39}$$

4.5 Derivation of $\frac{\partial E}{\partial w_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{ji}}$ as:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} \tag{40}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{41}$$

From key equations, we know that $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y}{\partial \phi_j} = w_{jk} \tag{42}$$

The equation $\frac{\partial \phi_j}{\partial a_I}$ can be written as

$$\begin{aligned}\frac{\partial \phi_j}{\partial a_I} &= \frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial a_I} \\ &= \frac{-zj}{\sigma^2}\phi_j.\frac{-(a_I - c_{ji})}{z_j}\end{aligned} \tag{43}$$

From key equations, we know that $a_I = \sum\limits_{i=1}^{I} x_i w_{ji}$, so $\frac{\partial a_I}{\partial w_{ji}}$ will be:

$$\frac{\partial a_I}{\partial w_{ji}} = x_i \tag{44}$$

Using equations (41), (42), (43) & (44) we can rewrite equation (40) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{ji}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} \\ &= \frac{-(t_k - y_k)w_{kj}\phi_j(a_I - c_{ji})x_i}{\sigma^2}\end{aligned} \tag{45}$$

4.6 Derivation of $\frac{\partial E}{\partial \sigma_j}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial \sigma_j} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}]\frac{\partial \phi_j}{\partial \sigma_j} \tag{46}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{47}$$

$$\frac{\partial y_k}{\partial \phi_j} = w_{kj} \tag{48}$$

$$\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}} \tag{49}$$

$$\frac{\partial \phi_j}{\partial \sigma_j} = \frac{z_j^2 \phi_j}{\sigma_j^3} \tag{50}$$

Using equations (47), (48), (49) & (50) we can rewrite equation (46) as:

$$\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}].\frac{z_j^2 \phi_j}{\sigma_j^3} \tag{51}$$

4.7 Final Equations of $\frac{\partial E}{\partial c_{ji}}$, $\frac{\partial E}{\partial w_{kj}}$, $\frac{\partial E}{\partial w_{ji}}$, $\frac{\partial E}{\partial \sigma_j}$

- $\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_j - c_j)}{\sigma^2}$
- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k).\phi_{j]}$
- $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_I - c_{ji}) x_i}{\sigma^2}$
- $\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}].\frac{z_j^2 \phi_j}{\sigma_j^3}$

5 Model 5: RBFNN model with updates of weights and centres using input e^{-x}

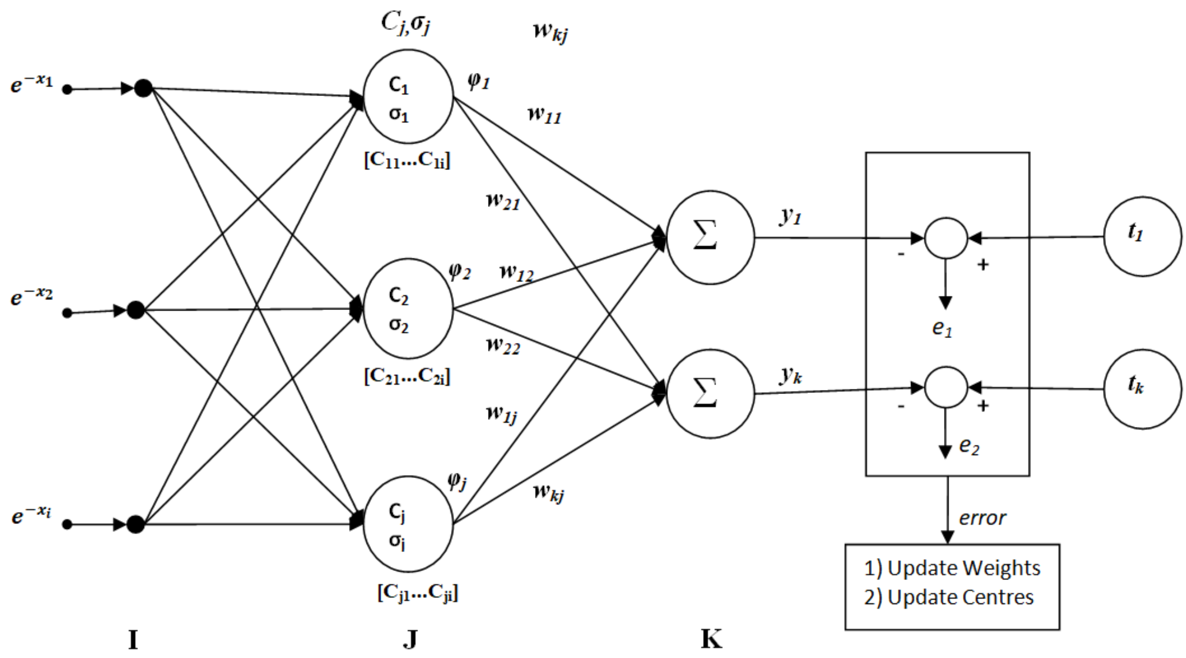

Figure 5: RBFNN model with updates of weights and centres using input e^{-x}

5.1 Key Equations

- $y_k = \sum\limits_j w_{kj}\phi_j$
- $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$
- $z_j = ||X - C_j|| = \sqrt{\sum\limits_j (e_i^{-x} - c_{ij})^2}$
- Cost function: $E = \frac{1}{2}\sum\limits_k (t_k - y_k)^2$

5.2 Gradient Descent Learning

- $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$
  where η_w = learning rate for update of weights
- $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$
  where η_c = learning rate for update of centres

5.3 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial w_{kj}} \tag{52}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum_k (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{53}$$

From key equations, we know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial w_{kj}}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial(\sum_j (w_{kj}\phi_j)}{\partial w_{kj}} \\ &= \phi_j\end{aligned} \tag{54}$$

Using equations (53) & (54) we can rewrite equation (52) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k)\phi_j\end{aligned} \tag{55}$$

5.4 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \tag{56}$$

From equation (2), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$
From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial \phi_j} &= \frac{\partial(\sum_j w_{kj}\phi_j}{\partial \phi_j} \\ &= w_{kj}\end{aligned} \tag{57}$$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial z_j}$ will be:

$$\begin{aligned}\frac{\partial \phi_j}{\partial z_j} &= \frac{\partial e^{\frac{-z^-2_j}{2\sigma_j^2}}}{\partial z_j} \\ &= -\frac{z_j \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^2} \\ &= -\frac{z_j \times \phi_j}{\sigma_j^2}\end{aligned} \tag{58}$$

From key equations, We know that $z_j = ||X - C_j|| = \sqrt{\sum_j (e_i^{-x} - c_{ij})^2}$, so $\frac{\partial z_j}{\partial c_{ji}}$ will be:

$$\begin{aligned}
\frac{\partial z_j}{\partial c_{ji}} &= \frac{\partial(\sqrt{\sum_i (e_i^{-x} - c_{ji})^2})}{\partial c_{ji}} \\
&= -\frac{2 \times (e_i^{-x} - c_{ji})}{2 \times \sqrt{\sum_i (e_i^{-x} - c_{ji})^2}} \\
&= -\frac{(e_i^{-x} - c_{ji})}{z_j}
\end{aligned} \tag{59}$$

Using equations (2), (57), (58) & (59) we can rewrite equation (56) as:

$$\begin{aligned}
\frac{\partial E}{\partial c_{ji}} &= [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} \\
&= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{-(z_j \times \phi_j)}{\sigma_j^2} \times \frac{-(e_i^{-x} - c_{ji})}{z_j} \\
&= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(e_i^{-x} - c_{ji})
\end{aligned} \tag{60}$$

5.5 Final equations of $\frac{\partial E}{\partial w_{kj}}$ $\frac{\partial E}{\partial c_{ji}}$

- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
- $\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} = [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(e_i^{-x} - c_{ji})$

6 Model 6 : RBFNN model with updates of weights in both layers, centres and standard deviations and using input e^{-x}

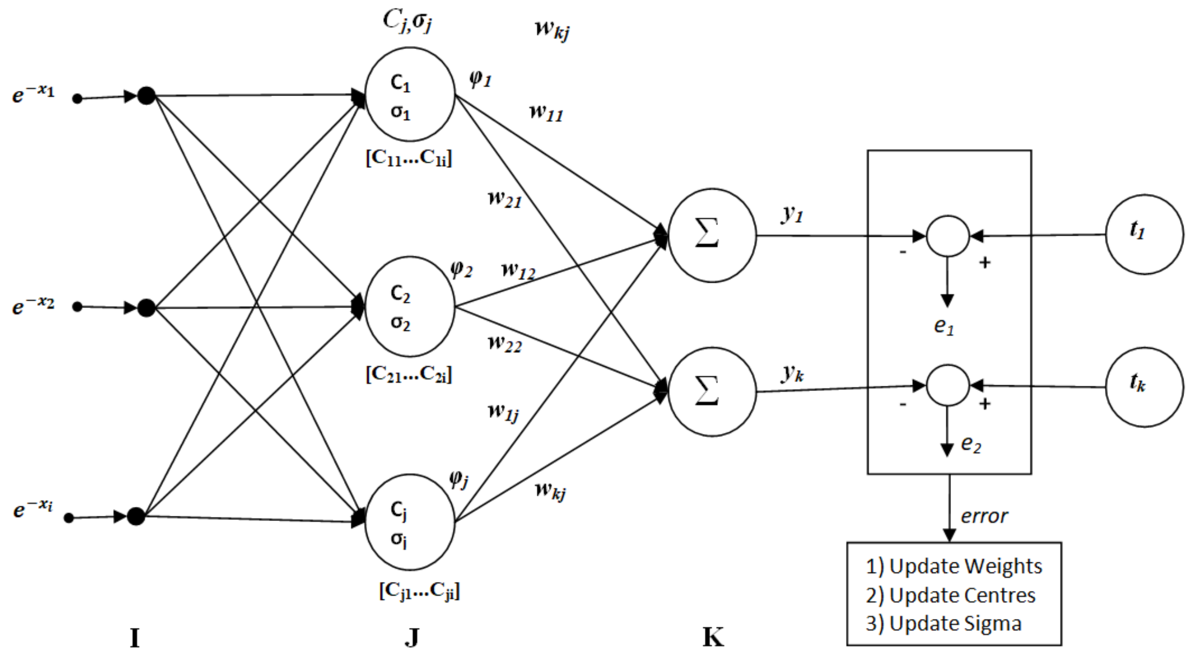

Figure 6: RBFNN model with updates of weights in both layers, centres and standard deviations and using input e^{-x}

6.1 Key Equations

- $a_I = \sum_{i=1}^{I} e_i^{-x} w_{ji}$
- $z_j = (||a_I - c_j||)$
  Let $||a_I - c_j|| = z =$ Euclidean distance
- $\phi(z) = e^{\frac{-z_j^2}{2\sigma^2}}$
- $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$
- Cost function: $E = \frac{1}{2} \sum (t_k - y_k)^2$

6.2 Gradient Descent Learning

- Learning of centres:

$$c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$$

  where $\eta_c =$ learning rate for update of centres

- Learning of weights:
  For w_{kj},
  $$w_{kj}(t+1) = w_{kj}(t) - \eta_{w_2}\frac{\partial E}{\partial w_{kj}}$$
  For w_{ji},
  $$w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2}\frac{\partial E}{\partial w_{ji}}$$
  where η_{w_1} and η_{w_2} = learning rate for update of weights
- Learning of sigma:
  $$\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma\frac{\partial E}{\partial \sigma_j}$$
  where η_σ = learning rate for update of sigmas

6.3 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \tag{61}$$

From key equations, we know that cost function, $E = \frac{1}{2}\sum(t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum(t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{62}$$

From key equations, we know that $y_k = \sum_{j=1}^{J}\phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y_k}{\partial \phi_j} = w_{jk} \tag{63}$$

The equation $\frac{\partial \phi_j}{\partial c_j}$ can be written as

$$\begin{aligned}\frac{\partial \phi_j}{\partial c_j} &= \frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_j} \\ &= \frac{-zj}{\sigma^2}\phi_j.\frac{-(a_j - c_j)}{z_j}\end{aligned} \tag{64}$$

Using equations (62), (63) & (64) we can rewrite equation (61) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{ji}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} \\ &= \frac{-(t_k - y_k)w_{kj}\phi_j(a_j - c_j)}{\sigma^2}\end{aligned} \tag{65}$$

6.4 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k).\phi_j\end{aligned} \tag{66}$$

6.5 Derivation of $\frac{\partial E}{\partial w_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{ji}}$ as:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial a_I} . \frac{\partial a_I}{\partial w_{ji}} \tag{67}$$

From key equations, we know that cost function, $E = \frac{1}{2} \sum (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial y_k} &= \frac{\partial (\frac{1}{2} \sum (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k) \end{aligned} \tag{68}$$

From key equations, we know that $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\frac{\partial y_k}{\partial \phi_j} = w_{jk} \tag{69}$$

The equation $\frac{\partial \phi_j}{\partial a_I}$ can be written as

$$\begin{aligned} \frac{\partial \phi_j}{\partial a_I} &= \frac{\partial \phi_j}{\partial z_j} . \frac{\partial z_j}{\partial a_I} \\ &= \frac{-zj}{\sigma^2} \phi_j . \frac{-(a_I - c_{ji})}{z_j} \times -e^{-x} \end{aligned} \tag{70}$$

From key equations, we know that $a_I = \sum_{i=1}^{I} x_i w_{ji}$, so $\frac{\partial a_I}{\partial w_{ji}}$ will be:

$$\frac{\partial a_I}{\partial w_{ji}} = e_i^{-x} \tag{71}$$

Using equations (68), (69), (70) & (71) we can rewrite equation (67) as:

$$\begin{aligned} \frac{\partial E}{\partial w_{ji}} &= \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial a_I} . \frac{\partial a_I}{\partial w_{ji}} \\ &= \frac{(t_k - y_k) w_{kj} \phi_j (a_I - c_{ji}) e_i^{-2x}}{\sigma^2} \end{aligned} \tag{72}$$

6.6 Derivation of $\frac{\partial E}{\partial \sigma_j}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial \sigma_j} = [\sum_k \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j}] \frac{\partial \phi_j}{\partial \sigma_j} \tag{73}$$

From key equations, we know that cost function, $E = \frac{1}{2} \sum (y^d - y)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\frac{\partial E}{\partial y_k} = \frac{\partial (\frac{1}{2} \sum (y^d - y)^2)}{\partial y_k} == -(t_k - y_k) \tag{74}$$

$$\frac{\partial y_k}{\partial \phi_j} = w_{kj} \tag{75}$$

$$\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}} \tag{76}$$

$$\frac{\partial \phi_j}{\partial \sigma_j} = \frac{z_j^2 \phi_j}{\sigma_j^3} \tag{77}$$

Using equations (74), (75), (76) & (77) we can rewrite equation (73) as:

$$\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}].\frac{z_j^2 \phi_j}{\sigma_j^3} \tag{78}$$

6.7 Final Equations of $\frac{\partial E}{\partial c_{ji}}$, $\frac{\partial E}{\partial w_{kj}}$, $\frac{\partial E}{\partial w_{ji}}$, $\frac{\partial E}{\partial \sigma_j}$

- $\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_j - c_j)}{\sigma^2}$
- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k).\phi_j$
- $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} = \frac{(t_k - y_k) w_{kj} \phi_j (a_I - c_{ji}) e_i^{-2x}}{\sigma^2}$
- $\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}].\frac{z_j^2 \phi_j}{\sigma_j^3}$

7 Tables

Table 1: Models and Equations

Models	Equations
Model 1	**Key Equations:** $y_k = \sum_j w_{kj}\phi_j$ $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$ $z_j = \|\|X - C_j\|\| = \sqrt{\sum_j (x_i - c_{ij})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$ **Gradient Descent Learning:** $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$ $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$ **Final equations:** $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$ $\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}].\frac{\partial \phi_j}{\partial z_j}.\frac{\partial z_j}{\partial c_{ji}} = [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})$
Model 2	**Key Equations:** $y_k = \sum_j w_{kj}\phi_j$ $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$ $z_j = \|\|X - C_j\|\| = \sqrt{\sum_j (x_i - c_{ij})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$ **Gradient Descent Learning:** $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$ $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$ $\sigma_{ji}(t+1) = \sigma_{ji}(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_{ji}}$ **Final equations:** $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$ $\frac{\partial E}{\partial c_{ji}} = [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(x_i - c_{ji})$

	$\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - yk)w_{kj}]\frac{z_j^2 \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^3}$
Model 3	**Key Equations:** $a_I = \sum_{i=1}^{I} x_i w_{ji}$ $z_j = (\lvert\lvert a_I - c_j \rvert\rvert)$ Let $\lvert\lvert a_I - c_j \rvert\rvert = z =$ Euclidean distance $\phi_z = e^{\frac{-z_j^2}{2\sigma^2}}$ $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$ Cost function: $E = \frac{1}{2}\sum(t_k - y_k)^2$ **Gradient Descent Learning:** $c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$ $w_{kj}(t+1) = w_{kj}(t) - \eta_{w_1} \frac{\partial E}{\partial w_{kj}}$ $w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{ji}}$ **Final equations:** $\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k)w_{kj}\phi_j(a_j - c_j)}{\sigma^2}$ $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k).\phi_j$ $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k}.\frac{\partial y_k}{\partial \phi_j}.\frac{\partial \phi_j}{\partial a_I}.\frac{\partial a_I}{\partial w_{ji}} = \frac{-(t_k - y_k)w_{kj}\phi_j(a_I - c_{ji})x_i}{\sigma^2}$
Model 4	**Key Equations:** $a_I = \sum_{i=1}^{I} x_i w_{ji}$ $z_j = (\lvert\lvert a_I - c_j \rvert\rvert)$ Let $\lvert\lvert a_I - c_j \rvert\rvert = z =$ Euclidean distance $\phi(z) = e^{\frac{-z_j^2}{2\sigma^2}}$ $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$ Cost function: $E = \frac{1}{2}\sum(t_k - y_k)^2$ **Gradient Descent Learning:** $c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$ $w_{kj}(t+1) = w_{kj}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{kj}}$ $w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{ji}}$ $\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$ **Final equations:**

	$\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_j - c_j)}{\sigma^2}$ $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k) . \phi_{j]}$ $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial a_I} . \frac{\partial a_I}{\partial w_{ji}} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_I - c_{ji}) x_i}{\sigma^2}$ $\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}] . \frac{z_j^2 \phi_j}{\sigma_j^3}$
Model 5	**Key Equations:** $y_k = \sum_j w_{kj} \phi_j$ $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$ $z_j = \|\|X - C_j\|\| = \sqrt{\sum_j (e_i^{-x} - c_{ij})^2}$ Cost function: $E = \frac{1}{2} \sum_k (t_k - y_k)^2$ **Gradient Descent Learning:** $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$ $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$ **Final equations:** $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k) \phi_j$ $\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j}] . \frac{\partial \phi_j}{\partial z_j} . \frac{\partial z_j}{\partial c_{ji}} = [\sum_k -(t_k - y_k) w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times -(e_i^{-x} - c_{ji})$
Model 6	**Key Equations:** $a_I = \sum_{i=1}^{I} e_i^{-x} w_{ji}$ $z_j = (\|\|a_I - c_j\|\|)$ Let $\|\|a_I - c_j\|\| = z =$ Euclidean distance $\phi(z) = e^{\frac{-z_j^2}{2\sigma^2}}$ $y_k = \sum_{j=1}^{J} \phi_j w_{kj}$ Cost function: $E = \frac{1}{2} \sum (t_k - y_k)^2$ **Gradient Descent Learning:** $c_{ij}(t+1) = c_{ij}(t) - \eta_c \frac{\partial E}{\partial c_{ij}}$ $w_{kj}(t+1) = w_{kj}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{kj}}$ $w_{ji}(t+1) = w_{ji}(t) - \eta_{w_2} \frac{\partial E}{\partial w_{ji}}$ $\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$ **Final equations:**

	$\frac{\partial E}{\partial c_{ji}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial c_j} = \frac{-(t_k - y_k) w_{kj} \phi_j (a_j - c_j)}{\sigma^2}$ $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k) . \phi_j$ $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_k} . \frac{\partial y_k}{\partial \phi_j} . \frac{\partial \phi_j}{\partial a_I} . \frac{\partial a_I}{\partial w_{ji}} = \frac{(t_k - y_k) w_{kj} \phi_j (a_I - c_{ji}) e_i^{-2x}}{\sigma^2}$ $\frac{\partial E}{\partial \sigma_j} = [\sum_k -(t_k - y_k) w_{kj}] . \frac{z_j^2 \phi_j}{\sigma_j^3}$

Table 2: Details of structure of various types of RBFNN models

Models	Center	Sigma	Weight Update in Layers	Input
Model 1	Update	Fixed	Input	x
Model 2	Update	Update	Input	x
Model 3	Update	Fixed	Input and Output	x
Model 4	Update	Update	Input and Output	x
Model 5	Update	Fixed	Input	e^{-x}
Model 6	Update	Update	Input and Output	e^{-x}

Table 3: Parameters used for simulation study of Model 1

	structure	muc	muow
Classification	2-2-1	.0004	.0003
Direct modelling	3-3-1	.00002	.00002
Inverse modelling	6-6-1	.0005	.0005

Table 4: Parameters used for simulation study of Model 2

	structure	muc	mus	muow
Classification	2-2-1	0.004	.0004	.003
Direct modelling	3-3-1	.00002	.00001	.00002
Inverse modelling	6-6-1	.005	.0001	.005

Table 5: Parameters used for simulation study of Model 3

	structure	muc	muiw	muow
Classification	2-2-1	0.004	.0005	.003
Direct modelling	3-3-1	.00005	.001	.00002
Inverse modelling	6-6-1	.005	.0005	.005

Table 6: Parameters used for simulation study of Model 4

	structure	muc	mus	muiw	muow
Classification	2-2-1	0.004	.0004	.0005	.003
Direct modelling	3-3-1	.00005	.0005	.001	.00002
Inverse modelling	6-6-1	.005	.0001	.0005	.005

Table 7: Parameters used for simulation study of Model 5

	structure	muc	muow
Classification	2-2-1	0.0007	.0006
Direct modelling	3-3-1	.00005	.00005
Inverse modelling	6-6-1	.005	.005

Table 8: Parameters used for simulation study of Model 6

	structure	muc	mus	muiw	muow
Classification	2-2-1	0.0008	.0001	.0005	.0006
Direct modelling	3-3-1	.00005	.0002	.0005	.00005
Inverse modelling	6-6-1	.005	.0001	.0005	.005

7.1 XOR Classification

The Ex-OR gate can be viewed as a classifier to which four possible inputs are applied, and two possible outputs are obtained. When the inputs are identical, the output shows class 1, and when the inputs are dissimilar, the output shows another class.
The XOR classification using rbf algorithm is done to check which form model of rbf is best suited for higher accuracy and speed.
For the above data rbf model 6(which contains both sided weights and centre and sigma update) is the best model for both speed and accuracy.

Table 9: System specification of simulated problems

Problem	**Type of Problem**	**Linear Part**	**Non-linear Part**
Example 1	Classification	-	-
Example 2	System Identification (Direct Modelling)	$0.304z + 0.903z^{-1} + 0.304z^{-2}$	$\tan(a(k))$
Example 3	Channel Equalization (Inverse Modelling)	$0.304z + 0.903z^{-1} + 0.304z^{-2}$	$a(k) + 0.2a^2(k) - 0.1a^3(k)$

Table 10: Comparison of performance for classification problem (Ex. 1)

Models	RMSE	MAD	Iterations to Converge
Model 1	0.2078	0.2648	3960
Model 2	0.0508	0.1590	190
Model 3	0.0020	0.0921	130
Model 4	0.0017	0.1362	130
Model 5	0.0581	0.8903	3110
Model 6	0.0016	0.03033	70

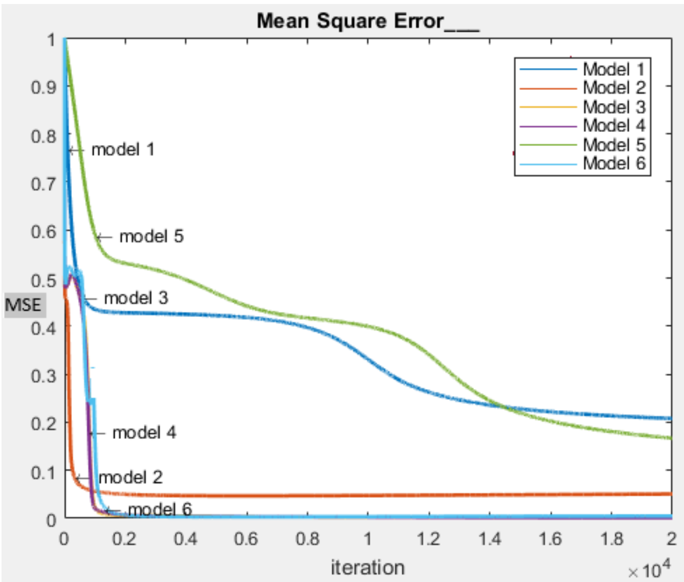

Figure 7: Comparison of convergence characteristics of various models in classification problem (Ex. 1)

7.2 System Identification

The system consists of linear and nonlinear parts.
Linear part:
$0.304z + 0.903z^{-1} + 0.304z^{-2}$
Non-linear part:
$\tan(a(k))$
(where a(k) is the output of the linear channel)
The input applied is a zero mean random input, and the output of the nonlinear system is mixed with additive white Gaussian noise (AWGN) with SNR of 30 db.
The same input set is applied to all RBF model. The error between the system output and the model is used to train the weights until convergence. After the training phase, the responses of the models and the known system are compared through which the simulation study and results are obtained.

Table 11: Comparison of performance of different models of non linear system identification problem (Ex. 2)

Direct Modelling

Models	RMSE	MAD	Iterations to Converge
Model 1	0.2308	1.8910	3452
Model 2	0.0524	1.7903	1791
Model 3	0.0153	0.5488	722
Model 4	0.0081	0.3834	689
Model 5	0.1750	0.2067	2085
Model 6	0.0040	0.0538	42

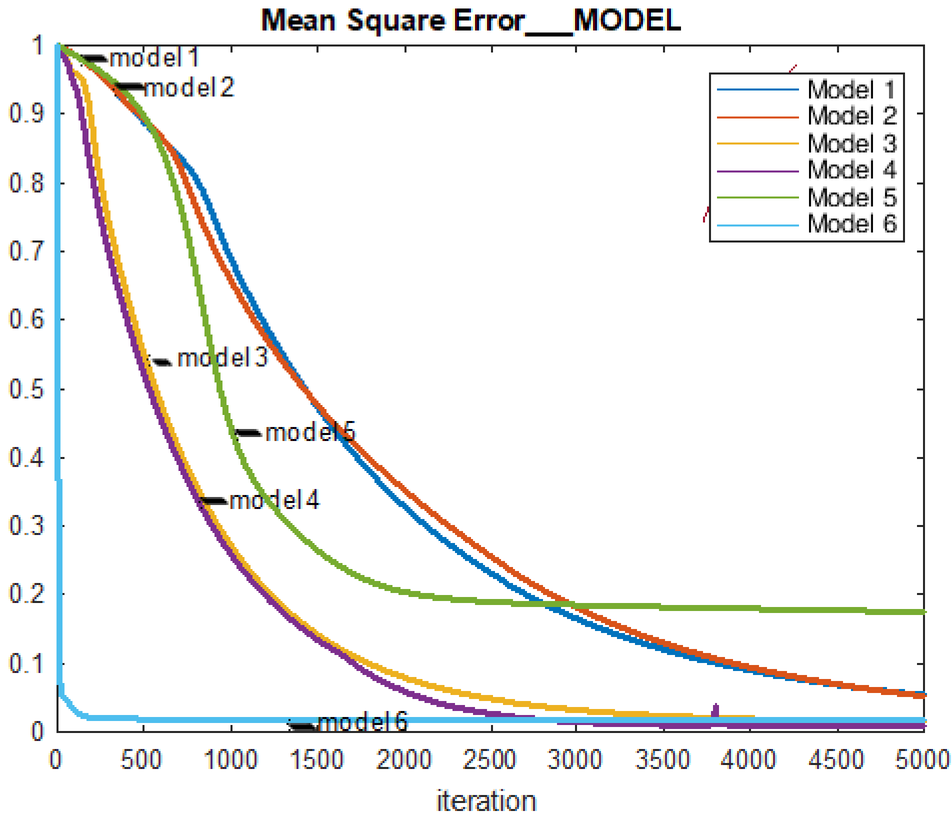

Figure 8: Comparison of convergence characteristics of different models on system identification (Ex. 2)

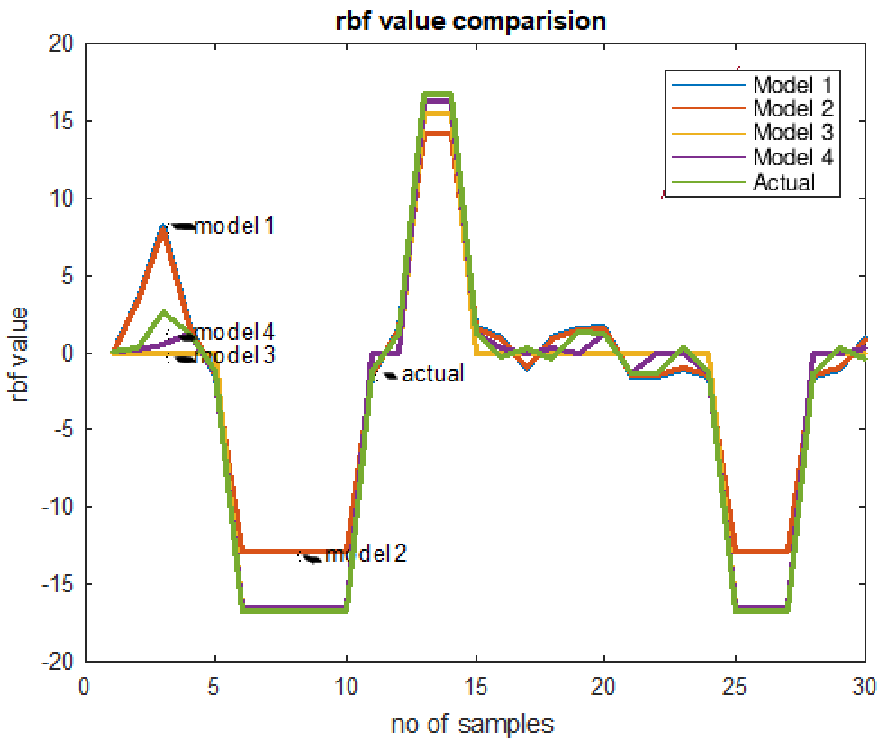

Figure 9: Comparison of output response of different models on non-linear system identification (Ex. 2)

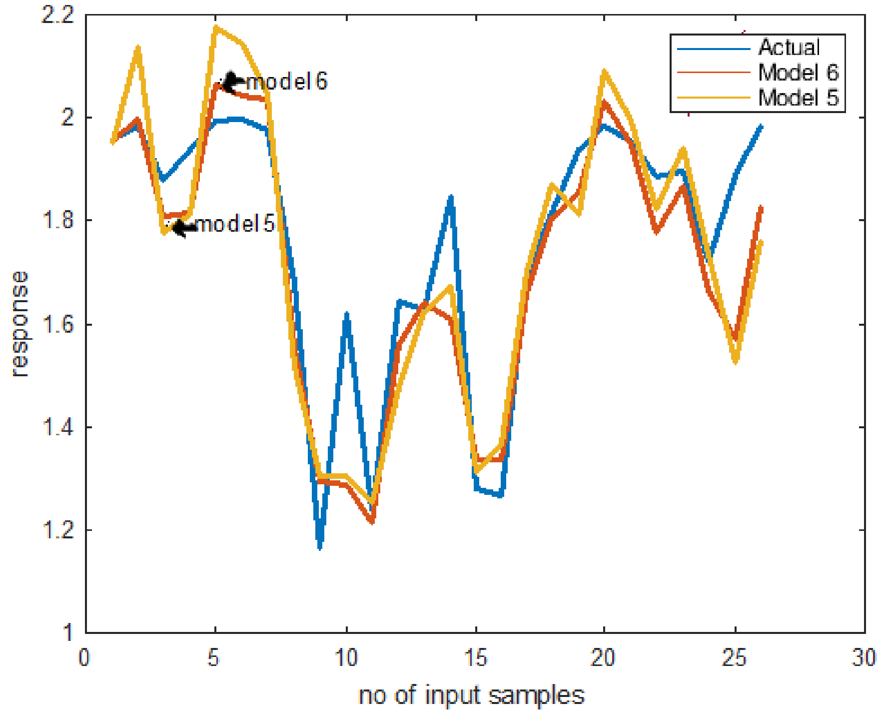

Figure 10: Comparison of output response of different models on non-linear system identification with input e^{-x} (Ex. 2)

7.3 Channel Equalization

In this modelling, zero mean random input is applied to a nonlinear digital channel(consisting of linear and non-linear parts) followed by AWGN of SNR 30 db. The output so obtained is distorted due to transmission through the channel plus additive noise.
The order of the channel is at least two times the order of the nonlinear channel, so there are 6 inputs for the RBF models. The output of the model is compared with the delayed version of the transmitted random input.
Linear part:
$0.304z + 0.903z^{-1} + 0.304z^{-2}$
Non-linear part:
$a(k) + 0.2a^2(k) - 0.1a^3(k)$
(where a(k) is the output of the linear channel)

Table 12: Comparison of performance of different models of non linear channel equalization problem (Ex. 2)

Inverse Modelling		
Models	RMSE	Iterations to Converge
Model 1	0.2176	394
Model 2	0.0199	59
Model 3	0.0083	36
Model 4	0.0081	31
Model 5	0.1519	214
Model 6	0.0076	20

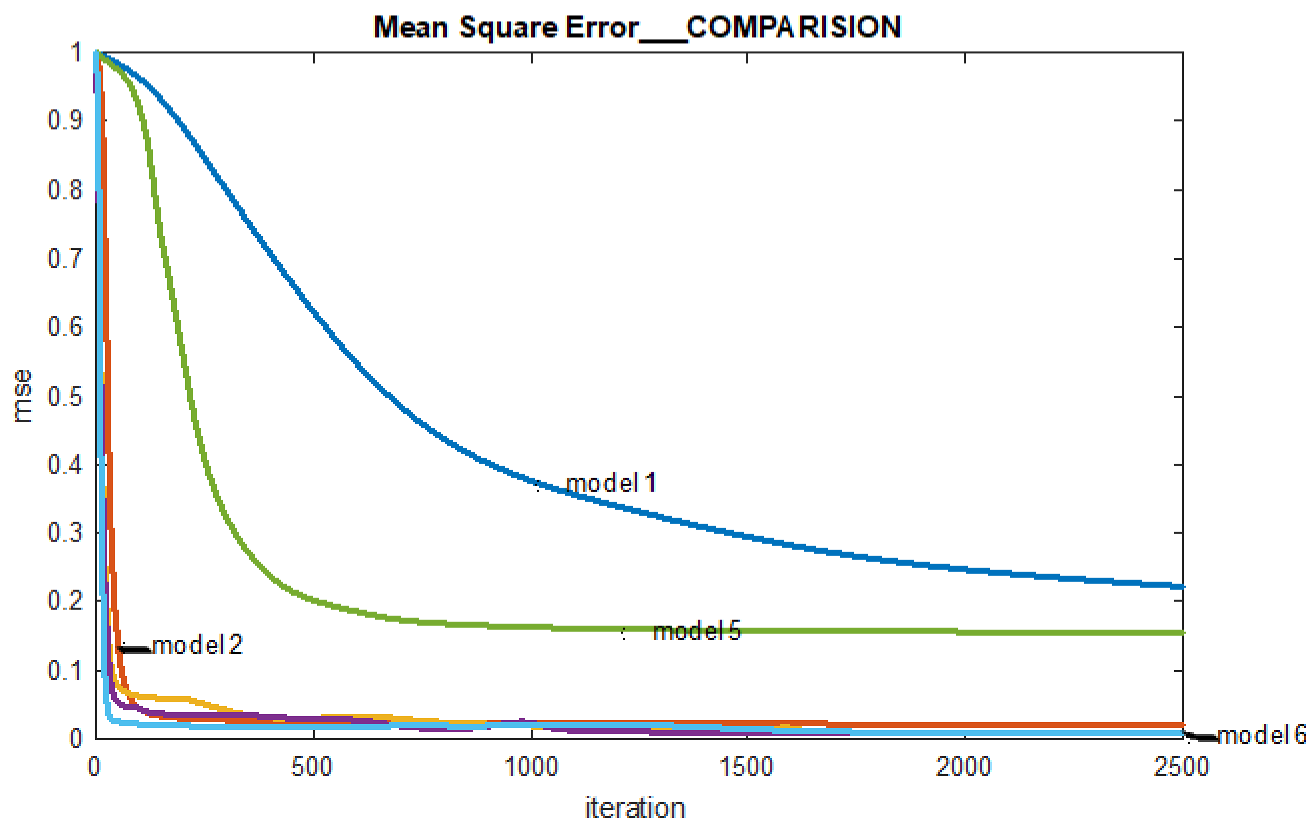

Figure 11: Comparison of convergence characteristics of different models on channel equalization (Ex. 3)

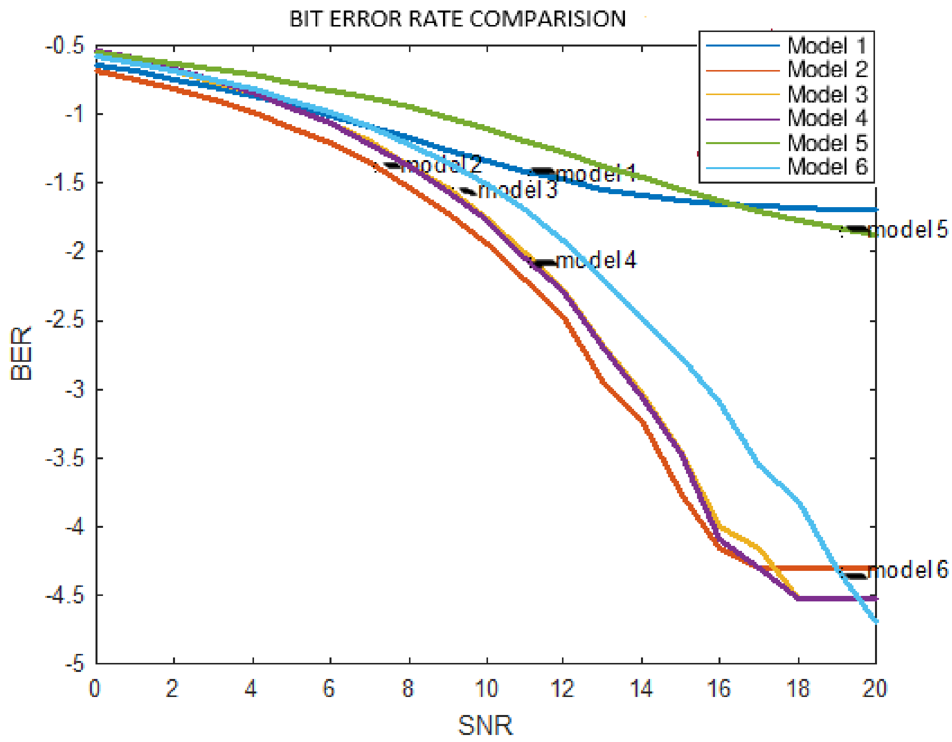

Figure 12: Comparison of BER of different models for non-linear channel equalization problem (Ex. 3)