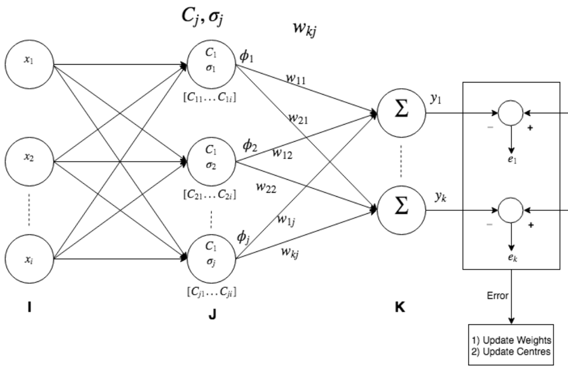
Analysis of Radial Basis Function Neural Networks (RBFNN)

Tanmay Khandait and Viral Patel

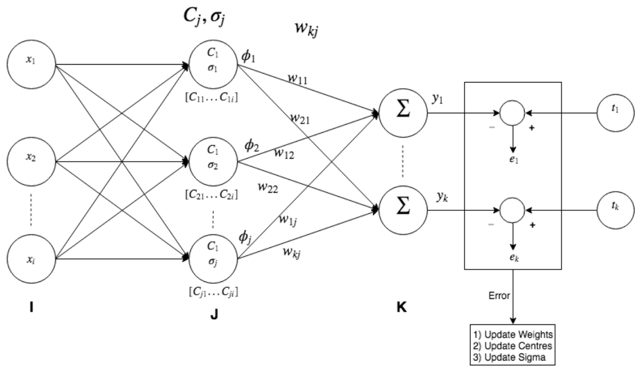
IIIT Vadodara

Abstract. In the following document, we have derived and analyzed variants of Radial Basis Function Neural Networks (RBFNN). There are four models we have worked on. The first part of the document consists of network diagrams, and a table of key equations and derivations of the back-propagation equations for all variant models.
The second part of the document contains test results of the rbfnn and its variants for X-OR classification problem.

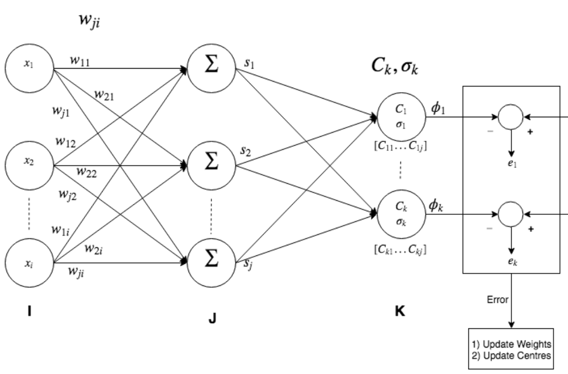
1 Radial Basis Function Neural Networks



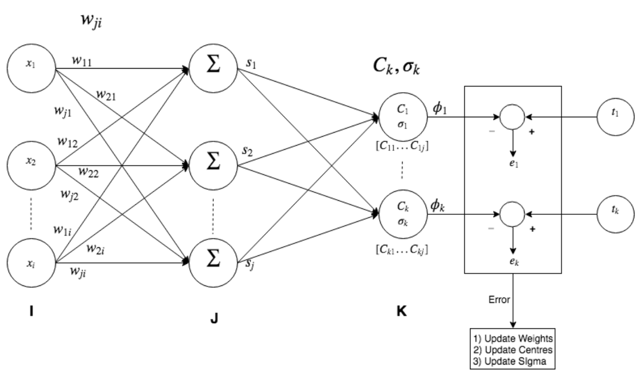
(a) Model 1



(b) Model 2



(c) Model 3



(d) Model 4

Table 1: Key equation, and back-propagation derivation results of all RBFNN models

	Key Equations	**Gradient Descent Learning**	**Computation of $\frac{\partial E}{\partial w}$, $\frac{\partial E}{\partial c}$ & $\frac{\partial E}{\partial \sigma}$**
	$y_k = \sum_j w_{kj}\phi_j$	Learning of weights:	$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
Model-1	$\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$	$w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$	
	$z_j = \|\|X - C_j\|\| = \sqrt{\sum_i (x_i - c_{ji})^2}$	Learning of centers:	$\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}}$
	Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$	$c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$	$= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$
		Learning of weights:	$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
	$y_k = \sum_j w_{kj}\phi_j$	$w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$	
Model-2	$\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$	Learning of centers:	$\frac{\partial E}{\partial c_{ji}} = [\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}}$
	$z_j = \|\|X - C_j\|\| = \sqrt{\sum_i (x_i - c_{ji})^2}$	$c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$	$= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$
	Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$	Learning of sigma:	$\frac{\partial E}{\partial \sigma_j} = [\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}] \cdot \frac{\partial \phi_j}{\partial \sigma_j}$
		$\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$	$= [\sum_k -(t_k - y_k)w_{kj}] \times \frac{\phi_j \times z_j^2}{\sigma_j^3}$
	$s_j = \sum_i w_{ji}x_i$	Learning of weights:	$\frac{\partial E}{\partial w_{ji}} = [\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}] \cdot \frac{\partial s_j}{\partial w_{ji}}$
Model-3	$\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$	$w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$	$= [\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})] \times x_i$
	$z_k = \|\|S - C_k\|\| = \sqrt{\sum_i (s_j - c_{kj})^2}$	Learning of centers:	$\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}}$
	Cost function: $E = \frac{1}{2}\sum_k (t_k - \phi_k)^2$	$c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$	$= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$
		Learning of weights:	$\frac{\partial E}{\partial w_{ji}} = [\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}] \cdot \frac{\partial s_j}{\partial w_{ji}}$
	$s_j = \sum_i w_{ji}x_i$	$w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$	$= [\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})] \times x_i$
Model-4	$\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$	Learning of centers:	$\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}}$
	$z_k = \|\|S - C_k\|\| = \sqrt{\sum_i (s_j - c_{kj})^2}$	$c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$	$= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$
	Cost function: $E = \frac{1}{2}\sum_k (t_k - \phi_k)^2$	Learning of sigma:	$\frac{\partial E}{\partial \sigma_k} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial \sigma_k}$
		$\sigma_k(t+1) = \sigma_k(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_k}$	$= -(t_k - \phi_k) \times \frac{\phi_k \times z_k^2}{\sigma_k^3}$

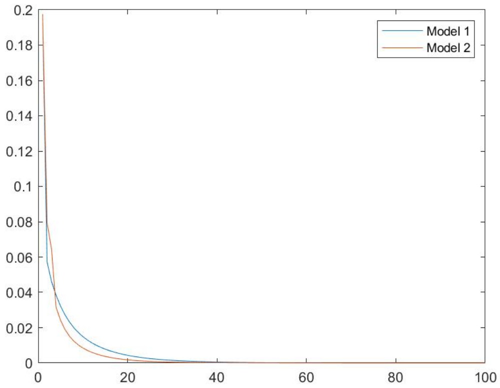
2 Training Results for XOR Problem

Table 2: RBF Network used and their respective learning parameters

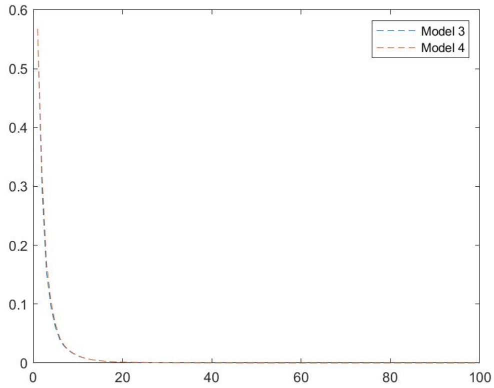
	Network configuration	**Learning Parameters**		
		Eta-Center	**Eta-Weight**	**Eta-Sigma**
		η_c	η_w	η_σ
Model-1	2-4-1	0.9099	0.4788	-
Model-2	2-4-1	0.9099	0.4788	0.5
Model-3	2-4-1	0.6	0.99	-
Model-4	2-4-1	0.4	0.99	10^{-3}

Table 3: Convergence Characteristics

	Number of iterations for convergence	**MSE at convergence**
Model-1	34	9.9083×10^{-4}
Model-2	25	9.1492×10^{-4}
Model-3	24	8.676×10^{-4}
Model-4	24	8.654×10^{-4}

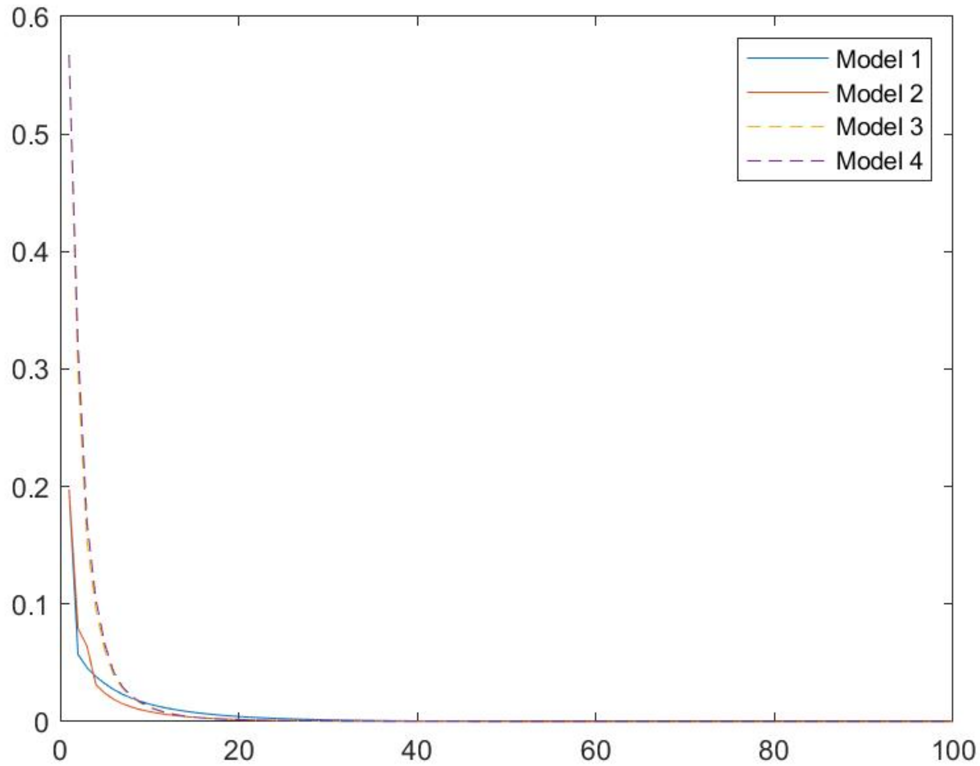


(e) Comparison of Model 1 and Model 2



(f) Comparison of Model 3 and Model 4

Fig. 1: Convergence characteristics comparison of all Models



3 XOR classification results

Table 4: Training results for RBFNN

	Testing Data		**Actual Output (AO)**	**Model Output (MO)**	**Deviation**	**Mean deviation**
Model-1	0.1	0.1	0.1	0.1034	0.0034	0.0019
	0.1	0.9	0.9	0.8976	0.0024	
	0.9	0.1	0.9	0.8986	0.0014	
	0.9	0.9	0.1	0.1006	0.0006	
Model-2	0.1	0.1	0.1	0.1041	0.0041	0.0021
	0.1	0.9	0.9	0.8990	0.0010	
	0.9	0.1	0.9	0.8989	0.0011	
	0.9	0.9	0.1	0.0979	0.0021	
Model-3	0.1	0.1	0.1	0.1026	0.0026	0.0011
	0.1	0.9	0.9	0.9000	0.0000	
	0.9	0.1	0.9	0.9000	0.0000	
	0.9	0.9	0.1	0.1017	0.0017	
Model-4	0.1	0.1	0.1	0.1029	0.0029	0.0012
	0.1	0.9	0.9	0.9000	0.0000	
	0.9	0.1	0.9	0.9000	0.0000	
	0.9	0.9	0.1	0.1018	0.0018	