Radial Basis Function (RBF) Neural Networks

Based on material provided by Professor W. Pedrycz

Neural Networks: Gradient-based Learning

- Updating weights relies on gradient of performance index

\[
\text{Weight}(\text{iter } + 1) = \text{Weight}(\text{iter}) - \alpha \nabla_{\text{weights}} Q
\]

- Speed of learning and convergence
- Local or global minimum
- Robustness of learning
- Dimensionality
- Computational overhead

Neural Networks: Overfitting and Generalization

Radial Basis Function (RBF) Neural Networks

Radial Basis Functions

Hidden neurons
Radial Basis Function (RBF): Basic classes
- Gaussian, triangular,...
- Wavelets
- Fourier-like (sin and cos)
- ....

RBF: Example
Euclidean metric
\[ v_i = \sqrt{(X - C_i)^2} \]

Ci = center vector of ith hidden layer neuron
X = input vector

Radial Basis Function (RBF): Main Properties

- Nonlinear transformation of data (normalization)
- Data patching - capturing domain knowledge

Linear vs. Nonlinear Normalization
Linear vs. Nonlinear Normalization

Receptive Fields: Data Patching

Receptive Fields: Granularity Issue

RBF Neural Networks: Learning Algorithm

Notation

\[
\text{target} = \begin{bmatrix}
\text{target}_1 \\
\text{target}_2 \\
\vdots \\
\text{target}_N
\end{bmatrix}
\]

\[
Z = \begin{bmatrix}
z_1 \\
z_2 \\
\vdots \\
z_N
\end{bmatrix}
\]

\[
\sum \begin{bmatrix}
w_1 \\
w_2 \\
\vdots \\
w_N
\end{bmatrix} \begin{bmatrix}
z_1 \\
z_2 \\
\vdots \\
z_N
\end{bmatrix} = \begin{bmatrix}
t_1 \\
t_2 \\
\vdots \\
t_N
\end{bmatrix}
\]
RBF Neural Networks: Learning Algorithm

Performance index

\[ Q = \sum_{k=1}^{N} [w^T z_k - \text{target}_k]^2 = [Zw - \text{target}]^T [Zw - \text{target}] \]

Minimization problem

\[ \frac{\partial Q}{\partial w} = 0 \]

Solution

\[ w = (Z^TZ)^{-1}Z^T \text{target} \]

Neural Networks in Data Mining

- Simple and interpretable architectures
- Easily scalable learning ability

Exploitation of Linguistic Blueprints

- Direct use of associations between linguistic granules
- Refinement of associations
- Linguistic learning

Network Representation of Information Granules
Direct use of Associations between Linguistic Granules

Computing activation levels of induced clusters
Summation of activation levels by a fuzzy neuron (fuzzy connections - contexts)
Output is a fuzzy set rather than a single number

\[ z_i = \left| \left| x - v_i \right| \right| \left| \left| x - v_j \right| \right| ^{2m-1} \]

\[ m_{ij} \] : all induced clusters

\[ Y = \sum_{i} W_i \otimes \{ z_i \} \]

\[ \Sigma \]

Refinement of Associations

Only modal values of contexts (fuzzy sets) used; their values are adjusted using standard gradient based method (RBF neural network)
Neural network

- huge and highly dimensional dataset
- large neural networks
- overfitting
- long, inefficient learning

Neural network

- network operates on the parameters of the linguistic granules
- limited number of granules
- limited number of parameters