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

RBF: Example

Euclidean metric

\[ v_i = \sqrt{(X - C_i)^2} \]

\( C_i \) = center vector of \( i \)th 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

Receptive Fields: Data Patching

Receptive Fields: Granularity Issue

RBF Neural Networks: Learning Algorithm

Notation:

\[
\text{target} = \begin{bmatrix}
\text{target}_1 \\
\text{target}_2 \\
\vdots \\
\text{target}_N
\end{bmatrix},
\quad
Z = \begin{bmatrix}
\mathbf{A}_1 \\
\mathbf{A}_2 \\
\vdots \\
\mathbf{A}_N
\end{bmatrix}
\]
RBF Neural Networks: Learning Algorithm

Performance index

\[ Q = \sum_{k=1}^{N} (w^T z_k - \text{target})^2 = [Z w - \text{target}]^T [Z w - \text{target}] \]

Minimization problem

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

Solution

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

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

\[ z_i = \frac{1}{\sum_{j: \text{all \ induced clusters}} (x - v_i)^2 m - 1} \]

Output is a fuzzy set rather than a single number

Direct use of Associations between Linguistic Granules

Computing activation levels of induced clusters

Normalization of activation levels by a fuzzy neuron (fuzzy connections - contexts)
Direct use of Associations between Linguistic Granules

\[ Y = \sum_{i} W_i \odot \{x_i\} \]

Refinement of Associations

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

Linguistic Learning

- Network operates on the parameters of the linguistic granules
- Limited number of granules
- Limited number of parameters

Linguistic Learning

- Huge and highly dimensional datasets
- Large neural networks
- Overfitting
- Long, inefficient learning