A Study on Pruning Methods for TAM Network

Isao Hayashi *, James R. Williamson **
* Hannan University
5-4-33, Amami-higashi, Matsubara, Osaka 580-8502, Japan, email: ihaya@hannan-u.ac.jp
** Into Networks
150 Cambridgepark Drive, Cambridge, MA 02140, email: jamesw@intonet.com

The TAM network[1] (Topographic Attentive Mapping network) based on a biologically-motivated neural network model is especially effective. The structure of the TAM network is shown in Figure 1. If the network's output prediction is not similar enough to the supervised output, the "attention" in the vigilance mechanism is invoked and the feedback parameter, ρ, is raised until either the prediction is satisfied or until the maximal vigilance level is reached. When ρ reached the maximal value, one category node is added to the category layer, and the weights of the category and the output layers are adjusted to be more correct.

In order to reduce an error between the supervised output and the network's output, and to acquire the rules from the network[2], a pruning mechanism, using information entropy, is introduced. The information entropy is defined as $H(i) = -\sum_{j=1}^{N} g_j \sum_{k=1}^{U} G_{jk} \log_2 G_{jk}$, where g_j is the degree of activation value of the j^{th} category in all of the categories and G_{jk} is the cumulative degree of activation value of data $s \in \psi_k$ in data set D. In the pruning algorithm, an importance of each feature is estimated. The necessity of link connections is simultaneously calculated according to the following three rules. Some connections to be unnecessary are removed.

Pruning Rule 1 If the first pruning condition, $G_{jk} \geq \eta$, is satisfied at each category j, the link connection between category j and class k is removed, where η is a threshold. Simultaneously, the connection between category j and feature i is removed.

Pruning Rule 2 If the second pruning condition, $\frac{1}{R}\sum_{s=1}^{R} \gamma_{js} < \theta$, is satisfied at each category j, the link connection between category j and feature i is removed, where θ is a threshold.

Pruning Rule 3 If the third pruning conditonon, $\varphi_{jK} \geq \xi$, is satisfied at class k, the link connection between class k and category j is removed, where ξ is a threshold and φ_{jK} is the cumulative degree of activation value of the data with correct class at the j^{th} category in all categories.

We should note that the pruning algorithm is a kind of fuzzy tuning method which adjusts the number of the features, classes and fuzzy rules.

The usefulness of the pruning algorithm is shown in a simple pattern problem, that the data of two groups are distributed at the two dimensions. After the learning, the correctness for training data is obtained as 97.8%. The correctness for checking data and the number of nodes in the network without pruning are respectively 82.2% and 2-4-2, which represents the number of nodes at the feature, category and output layers. The correctness and the node's number of the network after pruning are 97.8% and 1-2-2. Three fuzzy rules are acquired as shown in Figure 1.

[1] J.R.Williamson, *Technical Report CAS/CNS TR-99-027* (1999)
[2] I.Hayashi, *Proc. of 1996 Australian New Zealand Conference on Intelligent Information Systems*, pp.187-190 (1996)

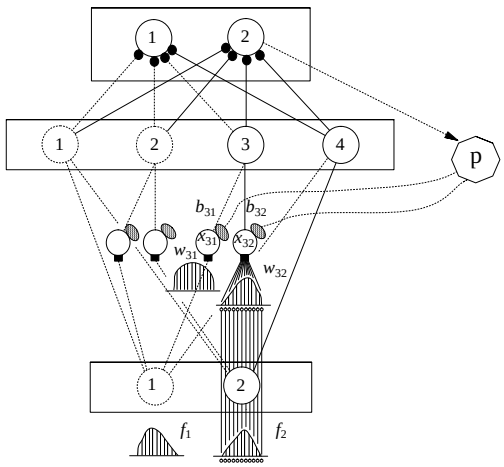

Figure 1: TAM Network after Pruning