Neural Network Ensemble with Temperature Control

4L−2

Pitoyo Hartono Shuji Hashimoto
Department of Applied Physics, School of Science and Engineering, Waseda University

1. Introduction

Recently a lot of researches have been devoted to investigate the performance of Neural Network Ensemble[1][2]. The Neural Network Ensemble is usually constructed to increase the reliability of the network's performance for fault tolerance system.

Due to its redundant feature, a neural network ensemble usually shows better performance compared to the conventional single network, it is however trained to perform in a single environment(a set of input patterns which require specific outputs), which we consider to be costly.

In this study, we propose an ensemble of neural network composed of a number of Multilayer Perceptrons(MLP), each with a unique ability, which have a mechanism to automatically activate the most appropriate member for a given environment while inhibit the unsuitable ones by using only a bilateral control of MLP's temperature.

2. Configuration of Neural Network Ensemble

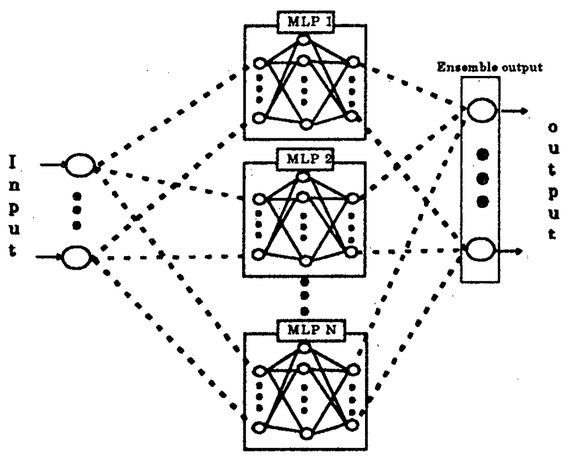

Figure 1. Neural Network Ensemble with N members

As shown in figure 1, the ensemble consists of a number of MLPs, where each of them runs independently of another. Inputs are simultaneously propagated to input units of each MLP. All of the outputs of the MLPs are connected to the corresponding units in the ensemble output layer with weights fixed to 1.

With our proposed algorithm only one member that is considered to be a specialist in the given environment will be activated, while the others will stay still.

3. Ensemble's Dynamic

In our proposed model the activation function of output neurons in i^{th} MLP is defined as:

$$O_m^{i,out}(u) = \frac{1}{1+\exp\left(-\frac{u}{T_i}\right)}, \quad u = \sum_{n=1}^{mid\_neuron} W_{nm}^{i,mid} O_{nm}^{i,mid} \quad (1)$$

Where $W_{nm}^{i,mid}$ is connection weight of m^{th} output neuron from n^{th} middle neuron, $O_n^{i,mid}$ is the output of n^{th} middle neuron in i^{th} MLP. T_i is the temperature of i^{th} MLP.

Regardless of the value of input u to the output neuron, the value of $O_m^{i,out}(u)$ always stays in the vicinity of 0.5 when T is big enough. For problems that require (0,1) binary value, 0.5 is an insignificant value.

The MLPs that are not relevant to the given environment can be deactivated by increasing its temperature. While MLP with a low temperature is permitted to actively run for a given environment, defining ensemble's k^{th} output O_k^{ens} as,

$$O_k^{ens} = \sum_{i=1}^{N} O_k^{i,out} - 0.5(N-1) \quad (2)$$

N is number of MLPs in the ensemble.

Furthermore, each MLP continuously correcting its connection weights while running in the given environment according to the Backpropagation learning rule[3], shown bellow.

$$W^i(t+1) = W^i(t) - \eta \frac{\partial E(t)}{\partial W^i(t)},$$

$$E = \frac{1}{2} \sum_{k=1}^{out\_neuron} (t_k - O_k^{i,out})^2 \quad (3)$$

W^i is weight vector of i^{th} MLP. E, t and $O^{i,out}$ are the error, teacher signal, and output of i^{th} MLP respectively, while η indicates the learning rate. For weights to output units, the correction value can be calculated as follow:

$$\frac{\partial E}{\partial w_{jk}^{i,mid}} = -\frac{1}{T_i} O_j^{i,mid}(t_k - O_k^{i,out})(1 - O_k^{i,out}) \quad (4)$$

$w_{jk}^{i,mid}$ indicates the weight from j^{th} middle neuron to k^{th} output neuron in the i^{th} MLP.

From (4) it is easy to see that if the temperature is large enough the correction value can be ignored, this is also true for connections in the other layers. It implies that, MLPs with high temperature will not only become inactive but will also become insensitive to learning process, which means that the expertise of a particular MLP to run in a particular environment will be preserved while the ensemble is running in a different environment.

4. Temperature Control

The idea of temperature control is to penalize particular ensemble members that performed badly in a given environment by increasing their temperature which will eventually inactivate them, and to recognize members which performed well by decreasing their temperature. Because eventually only one member should run actively for a given environment, a member that performs well will also increase others' temperature in order to dominate them. The opposite should be done when the performance is poor, in order to activate other members.

Temperature of i^{th} ensemble member T_i is calculated as follows,

$$T_i(t+F) = T_i(t) + p_i^{self}(M_i) + \sum_{j \neq i}^{N} p_j^{cross}(M_j) \quad (5)$$

$$p_i^{self}(M_i) = \alpha_1 H(M_i - 1) + \alpha_2 H(M_i - 0.75F) + \alpha_3$$

$$p_j^{cross}(M_j) = \beta_1 - \beta_2 H(M_j - 1) - \beta_3(M_j - 0.75F)$$

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & otherwise \end{cases}$$

F is temperature evaluation frequency, p^{self} and p^{cross} are self and cross penalty function respectively which calculate the penalty according to M_i, the number of misclassified patterns in i^{th} MLP observed in F trials, while α and β are temperature correction parameters.

The range of T is limited between 1 and 50.

5. Simulation

Ensemble consists of 3 members is used in this simulation. First 3 boolean function environments with 2 inputs and 1 output(XOR, NAND, NXOR) are presented to the ensemble sequentially. Because the ensemble does not have prior knowledge about the environments, each member is activated in turn to learn to became specialist of the respective environment(the learning process is shown in the first 3 curves of figure 2). Then, the environment is sequentially changed in NAND,NXOR,XOR order. The further curves show that, the ensemble can adapt very quickly to the changing environment, because temperature control always activates the most relevant member for any given environment. MSE in figure 2 indicates the ensemble's degree of unfitness to a given environment.

In this simulation the temperature parameters α and β are set to (16,4,-15) and (15,15,1) respectively.

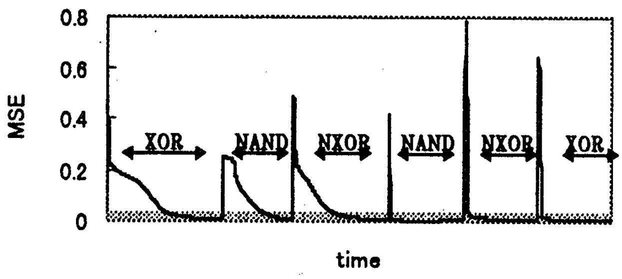

Figure2 Ensemble adaptation to the environment

Conclusion

In this study we have proposed a novel neural network ensemble model, that is able to run effectively in multiple environments, by activating the most appropriate member using bilateral temperature control. This ability is previously unachieved by conventional single MLP nor by formerly proposed ensemble models.

References

[1] Lars Kai Hansen: Neural Network Ensemble, IEEE Trans on Pattern Analysis and Machine Intelligence(1990), vol 12 no 10

[2] Michael Perrone,: When Networks Disagree: Ensemble Methods for Hybrid Neural Network, Neural Network for Speech and Image Processing(1993), Chapman-Hall

[3] D.E.Rumelhart et al: Parallel Distributed Processing(1986), vol.1 & 2, The MIT Press