The Comparison of Neural Network Learning Paradigms: Backpropagation, Simulated Annealing, Genetic Algorithm, and Tabu Search

Ming-Kuen Chen
Associate Professor, Industrial Engineering Department
National Taipei University of Technology, Taipei, Taiwan, R.O.C.

Abstract

Artificial neural networks (ANN) have successfully applied into various areas. But, How to effectively established network is the one of the critical problem. This study will focus on this problem and try to extensively study. Firstly, four different learning algorithms ANNs were constructed. The learning algorithms include backpropagation, simulated annealing, genetic algorithm, and tabu search. The experimental results of the above four different learning algorithms were tested by statistical analysis. The training RMS, training time, and testing RMS were used as the comparison criteria.

Keywords: Artificial Neural Network, Simulated Annealing, Genetic Algorithm, and Tabu Search

1. Introduction

As Bailey et al [1-2] surveyed, a series of applications has emerged to show that there are practical benefits to be reaped from a new technology, neural networks. The applications for this technology include diverse areas, such as pattern recognition, expert systems, decision support systems, neural language processing, neural modeling, information modeling of cognitive phenomena, and speech processing. Lately, neural networks have received more attention from academia. The researchers associated with of the machining process have been attracted by this technology so that several studies have been conducted, as reported in [3-8]. Besides, the applications in the machining process, their are several applications for neural networks in areas such as the cold-forging process [10]; failure diagnosis of servo-valve [11] and in modeling a markovian decision process [12].

Before applying the ANN, the critical problem is how to construct a network and this network must have the some characteristics like learning and the generality for the problem solving. Besides, how to effectively and fastly train an ANN and how to avoid the overtraining phenomena are the problems will be faced to the ANN researchers.

Simulated annealing, genetic algorithm, and tabu search are the optimal methodologies. It shows that these three optimal methodologies have the good performance in some applications [13-19]. This research try to systematically study the different ANN learning paradigms, like backpropagation, simulated annealing, genetic algorithm, and Tabu search and the experimental results will be compared.

2. Research Procedures and Data Collection

2.1 Research Procedures

The research procedures applied by this study are:

1) Collect the training and test data for the ANN.
2) Construct the ANNs for different ANN paradigms, like backpropagation, simulated annealing, genetic algorithm, and tabu search.
3) Training and test the above different ANN paradigms; and compare the results of training speed, test RMS, and overtraining RMS.
4) Statistical Analysis for the above experimental results from the different ANN paradigms.

In this study, the basic bachpropagation network topology is selected, shown in Figure 1. The sigmoid function is applied as the ANN transfer function; and each neuron all have threshold. During training phase, the different ANN paradigms were trained under maximum 2000 training cycles and the RMS is equal to 0.05.

2.2 Data Collection

For obtaining the data for training and testing the ANN, a ceramic creep feed grinding experimental was conducted. The ceramic material ZrO_2 was selected and ground by Jung JF 520 CNC Surface Grinder under down grinding mode. A 2^{5-1} fractional factorial experiment with two replicates, listed in Table 1, was designed for collecting the required data for the study. The factors included in this experimental design are three grinding wheel specifications: bond type, mesh size, and concentration, and two grinding process parameters: feed rate and depth of cut.

The grinding conditions used in this experiment are given in Table 2. The experiment was sequentially executed as follows:

1) The wheel were trued to ensure the concentricity of the wheel with a brake truing device and a silicon carbide truing wheel is used for this purpose.
2) The wheel was then dressed with an aluminum oxide stick to open up the wheel.
3) The wheel was stabilized by grinding two layers of alumina block about 0.12 inches.
4) The force and power data were collected while grinding the specimen.
5) The roughness of the ground workpiece was measured.
6) Finally, several profiles of the wheel were taken for counting the number of cutting edges on

the wheel surface.

7) The grinding force of x and y direction was measured by Norland Spectrum Analyzer; the horsepower required during grinding process was read from grinder amplitude meter.

8) The surface roughness of ground material was measured by Perthern S5P profilometer.

3. ANN Learning Paradigms

3.1 Backpropagation

The backpropagation ANN training algorithms are:

1) The weights between neurons and the threshold of each neuron were randomly produced.
2) Input one set input data into ANN and compute the output values; compute the deviation between actual output and ANN output.
3) Modified the weight by Delta rule.
4) Repeated step 2 until training number up to 2000 or the output RMS less than 0.05.

3.2 Simulated Annealing

The simulated annealing algorithms applied to ANN are:

1) The weights between neurons and the threshold of each neuron were randomly produced.
2) Input initial temperature.
3) Compute the ANN output. The energy function of the simulated annealing algorithm is replace by RMS. If RMS less than 0.05, or training cycle larger than 2000, stop the training phase. Otherwise, go to next step.
4) The weight updating
   (a) Randomize a weight updating, namely,

$$\Delta W_{ij}$$

   (b) Cooperate the weight change and old weight into a new weight.
5) (a) Compute the new ANN energy function, RMS_{new}.
   (b) If the $RMS_{new} \leq RMS_{old}$, new ANN network replace the old one, then go to Step 6. Check $RMS_{new} < 0.05$; If yes, stop training phase; Otherwise, randomize a random number, r,

$$r \in [0,1]$$

   (c) Compute the energy difference between the new and old ANN.
   (d) Compute the probability p,

$$p = e^{\frac{-\Delta E}{T}}$$

   (e) If r < p, then the new ANN replace the old ANN and turn down the temperature, then go to step 6; Otherwise, go to step 3.

6) Cumulative the number of training cycles. If the number of training cycle > 2000, then stop; Otherwise go to step 4.

3.3 Genetic Algorithm

The genetic algorithms applied to ANN are:

1) The weights between neurons and the threshold of each neuron of six ANN were randomly produced.

2. (a) Compute the outputs of six ANNs, and the RMS of the output results.

   (b) The fitness is defined by (1-RMS)

   (c) The relative fitness value is the fitness of individual ANN divided by the summation fitness of six ANN

(d) Follow the sequence of the relative value, namely, from ANN 1 to ANN 6, the coding value from 00-99 is cumulated.

2) Reproduction Period

   (a) Randomize six two-digit random numbers.

   (b) Referring the coding number from Step 2, determine the reproduction ANN by the value of random number.

3) Crossover Period

   (a) Six ANNs are divided into three-pair by their sequence;

   (b) Randomize three random numbers; these numbers determine the location of crossover; then, crossover three pairs ANN.

5) (a) Compute the outputs RMS from six new ANNs; Check if these RMS < 0.05; and training cycles > 2000 ?

   (b) If one of the above two conditions is satisfied, stop the training phase.

6) Mutation Period

   (a) Mutation probability is 0.08.

   (b) Randomly determine the mutation location

   (c) Cumulated the training cycles; go To Step 2.

3.4 Tabu Search

The algorithm of tabu search applied to ANN are:

1) The weights between neurons and the threshold of each neuron were randomly generated.

2) Let Tabu List ={}; CRMS (the current RMS)= 9999; AL (the aspiration level) = 0.1 ; TC (training cycle) = 0. BRMS (the best RMS) = 9999.

3) (a) Generate an ANN output and compute ANNRMS (the ANN output RMS).

   (b) Check if ANNRMS < CRMS ?

      If yes, check if this training pattern is in Tabu List or not ?

         If yes, check if ANNRMS < AL ?

If yes, then BRMS = ANNRMS; Otherwise go to step 3(a).

4) Check if does the stopping criteria satisfy ? (namely, BRMS =0.05 or TC = 2000.)
   If no, let AL=AL - 0.001; TC =TC + 1; and update Tabu List.
   Otherwise, stop the training phase.

4. Research Results and Discussion

4.1 Statistical Analysis

Based on the results of 2nd experimental grinding process, these four ANN learning paradigms were applied and the ANN outputs are normal grinding force, surface roughness, grinder power and specific heat. The outputs of ANN were tested by statistical hypotheses test by 5% significance level. In this study, t-paired tests were selected; null hypothesis: u1=u2, and the alternative hypothesis: u1=u2,. The t-value is compare with $t_{0.25}$ = +0.2131, in order to make a decision: whether can't reject or reject the null hypothesis.

The statistical results are shown in Table 3 to 6. From those tables, it shows that the outputs from different ANN learning paradigms only the grinder horsepower and specific energy learned by genetic algorithm and tabu search are rejected by null hypothesis; it means the output averages value of ANN have different.

4.2 ANN Training and Test

These four learning paradigms applied to ANN; the training processes were showed as Figure 2. There are some observations can be obtained:

1. It shows that the ANN error is reduced as learning cycle increased during the training process by the backpropogation learning paradigm.
2. It shows that the ANN error is reduced when training cycle increases until training cycle is 1800 during the training process by the simulated annealing learning paradigm.
3. It shows that he learning behavior of two learning paradigms, genetic algorithm and tabu search, looks similarly; namely, the ANN error have small vibration about the ANN value, 0.25.

5 Conclusions

The purpose of this research is to study the ANN learning behavior using four learning paradigms, backpropogation, simulated annealing, genetic algorithm, and tabu search. The results of these four learning paradigms were analyzed by statistical test.

In this research, the actual experimental data of the creep feed grinding process were applied to above four ANN learning paradigms. From the results of learning process, the simulated and baporgation have better performance than the other two.

Reference

1. Biley, D. Thomposon, D., and Feinstein, J., "The Practical Side of Neural Networks: Part I", PC AI, Nov/Dec., 1988, pp. 33-36.
2. Biley, D. Thomposon, D., and Feinstein, J., "The Practical Side of Neural Networks: Part II", PC AI, March/April, 1989, pp. 56-58.
3. Dornfeld, D.A., "Neural Network Sensor Fusion for Tool Condition Monitoring", Annals of CIRP, 39/1, 1990, pp. 101-105.
4. Rangwala, S. and Dornfeld, D.A.,"Integration of Sensors via Neural Networks for Detection of Tool Wear States", in Intelligent and Integrated Manufacturing Analysis and Synthethesis, ASME PED-25, 1987, pp. 109-120.
5. Rangwala, S. and Dornfeld, D.A., "Learning and Optimization of Machining Operation Using Computing Abilities of Neural Networks", IEEE Trans. On Systems, Man, and Cybernetic, Vol. 19, No. 2, March/April, 1989, pp. 299-234.
6. Rangwala, S. and Dornfeld, D.A., " Sensor Integration Using Neural Networks for Intelligent Tool Condition Monitoring", ASME J. of Engg,. for Industry, Vol. 112, Aug. 1990, pp. 219-228.
7. Okafor, A.C., Marcus, M. and Tipirneni, R., "Multiple Sensor Integration Via Neural Networks for Estimating Surface Roughness and Bore Tolerance in Circular End Milling", Trans of NAMRI/SME, 1990, pp. 128-136.
8. Tansel, I. N., "Neural Network Approach for Representation and Simulation of #D Cutting Dynamics", Trans. Of NAMRI/SME, 1990, 192-200.
9. Osakada, K., Yang, G.M., Nakamura, T., and Morio., "Expert System for Cold-Forging Process Based on FEM Simulation" Annul of the CIRP, 39/11, pp. 249-252.
10. Satri, T., English, J., and Krishnamurthi, M., "Modeling a Markovian Decision Process by Neural Network", Computers and Engineering, Vol.. 17, No. 1-4, 1989, pp. 464-468.
11. Yamashina, H., Kumamoto, H., Okumura, S., and Ikesaki, T., "Failure Diagnosis of a Servovalve by Neural Networks with New Learning Algorithm and Structure Analysis", Int. J. Prod. Res., Vol. 28, No. 6, 1990, pp. 1009-1021.
12. Ackley, D.H., Hiton, G.E., and Sejnowski, T.J., "A Learning Algorithm for Boltzmann Machines", Cognitive Science, No. 9, 1985, pp. 147-169.
13. Hiton, G.E., Sejnowski, T.J., "Learning and Relearning in Boltzmann Machines", in D.E. Rumehart and J.L. McClelland, Editors, Parallel Distributed Processing, Explorations in Microstructure of Cognition, MIT Press, Cambridge, MA, 1986, pp. 282-317.
14. Goldberg, D.E., Genetic Algorithm in Search, Optimization and Machine Learning, 1989, Reading Mass: Addison-Wesley.

15. Dodd, N., "Optimization of Network Structure using Genetic Techniques", IJCNN, 1990, Vol. III, pp. 965-970.
16. Harp, S.A. and Samad, T., "Genetic Synthesis of Neural Network Architecture", in Handbook of GA, edited by Davis, L., 1991, Van Nostrand Reiehold.
17. Montana, D. J., and Vavis, L., "Training Feedforward Neural Networks Using Genetic Algorithms," 11th Int. Joint Conf. On AI, 1989, pp. 762-767.
18. Miller, G.F. and Todd, P.M. ,"Designing Neural Networks using Genetic Algorithms", 3rd GA Conf., 1990, pp. 378-384.
19. Werra, D., and Hertz, A., "Tabu Search Techques", OR Spectrum, 1989, Vol. 11, pp. 131-141.
20. Beyer, D.A. and Ogier, R.G., "Tabu Learning: A Neural Network Search Method for Solving Nonconvex Optimization Problems", IJCNN, 1991, pp. 935-961, Singapore.

Table 1 Grinding Condition

Wheel Size : 7" x 0.25" x 0.125"

Wheel Speed: 4400 ft/min (2400 rpm)

Coolant : Cincinnati CIMTECH GL 2015, diluted in water in 1:50 proportion

Truing: Silicon Carbide truing with 0.001" increment

Dressing: Constant feeding of Al_2O_3 stick at speed of 1000 ft/min (545 Rpm)

Grinding Condition for wheel stabilization:

Depth of cut : 0.05"

Table Speed : 10 in/min

Table 2 2^{5-1} Fractional Factorial Experimental Design

Experimental Factors	Two-level
1. Wheel Bond Type	Resinoid, Vitrified
2. Wheel Mesh Size	80, 100
3. Wheel Concentration	50, 100
4. Feed Rate	6.8, 16.2 (inch/min)
5. Depth of Cut	0.058, 0.0102 (inch)

Table 3 Statistical Pair-test of Normal Grinding Force for Four ANN Learning Paradigms

	BP vs. SA	BP vs. GA	BP vs. TB	SA vs. GA	SA vs. TB	GA vs. TB
Average	3.999	0.5075	1.85	4.507	5.849	1.342
Std. Dev.	1.016	9.651	10.671	18.161	15.161	2.77
t-value	0.098	0.018	0.064	0.155	0.201	0.586

Table 4 Statistical Pair-test of Roughness for Four ANN Learning Paradigms

	BP vs. SA	BP vs. GA	BP vs. TB	SA vs. GA	SA vs. TB	GA vs. TB
Average	0.016	0.018	0.017	0.002	0.001	0.0005
Std. Dev.	0.034	0.048	0.047	0.014	0.001	0.001
t-value	1.274	1.462	1.42	0.441	0.304	1.192

Table 5 Statistical Pair-test of Grinder Horsepower for Four ANN Learning Paradigms

	BP vs. SA	BP vs. GA	BP vs. TB	SA vs. GA	SA vs. TB	GA vs. TB
Average	0.067	0.428	0.029	0.361	0.038	0.399
Std. Dev.	0.054	3.579	3.477	3.533	3.431	0.102
t-value	0.051	0.452	0.031	0.386	0.041	4.338

Table 6 Statistical Pair-test of Specific Energy for Four ANN Learning Paradigms

	BP vs. SA	BP vs. GA	BP vs. TB	SA vs. GA	SA vs. TB	GA vs. TB
Average	0.025	0.067	0.148	0.042	0.123	0.081
Std. Dev.	0.017	1.268	1.186	1.259	1.267	0.092
t-value	0.054	0.203	0.447	0.128	0.374	2.251

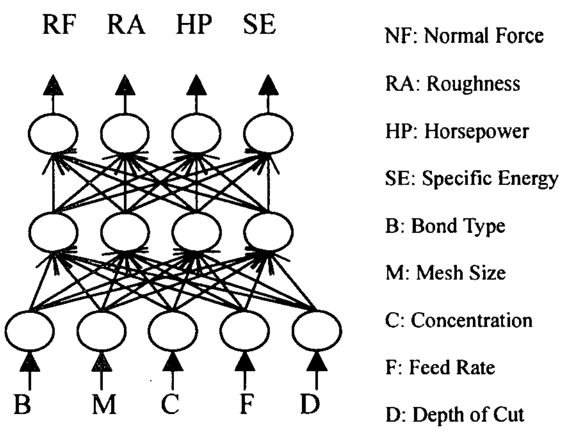

Figure 1 ANN Topology

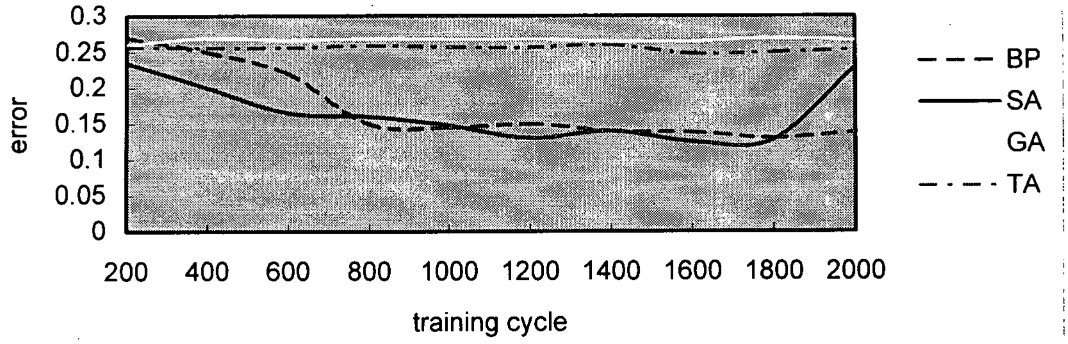

Figure 2 The Training Phase of Four Learning Paradigms