R.J. Marks II, "Neural networks for classification and regression",
Proc. of the First Workshop on Neural Networks: Academic/Industrial/NASA/Defense,
Auburn University and Conference Center, 4-6 February, 1990, Auburn, Alabama, pp.45-49

PROCEEDINGS OF THE

FIRST WORKSHOP ON NEURAL NETWORKS: ACADEMIC/INDUSTRIAL/NASA/DEFENSE

WNN-AIND 90

February 5-6, 1990

Auburn University Hotel and Conference Center, Auburn, Alabama

SPACE POWER INSTITUTE

NASA HEADQUARTERS

COSPONSORED

AUBURN UNIVERSITY RESEARCH EXTENSION INSTRUCTION FOR THE ADVANCEMENT OF SCIENCE AND ARTS

Neural Networks for Classification and Regression

*Robert J. Marks II**

Interactive Systems Design Laboratory
Department of Electrical Engineering, FT-10
University of Washington
Seattle, Washington 98195

Abstract

There have been a number of applications proposed for artificial neural network computational structures. The neural network architecture currently receiving most attention as a viable computational paradigm with competitive performance attributes is the layered perceptron. In this paper, we will address the relative performance of the layered perceptron and suggest query based techniques by which training can be improved.

Relative Performance

Do layered perceptrons perform better than other classifiers and regression machines? By comparison with some other high performance classifiers and regression machines, the current answer is yes - but not by much. Possibly there is an underlying limit of performance placed on all classifiers and regression machines that cutting edge algorithms are approaching. If so, then secondary performance attributes such as training speed and implementation ease must be addressed as primary.

Other artificial neural networks have fallen from favor in an application sense because, quite simply, they are not competitive with other more conventional approaches. The same question must be posed in regard to the layered perceptron. Does the layered perceptron preform better than other classifiers or regression machines programmed from examples using supervised learning? Although abstract analysis of this question may be possible in some cases, it must ultimately be answered in regard to actual data. Comparisons of the layered perceptron have been performed with *classification and regression trees* (CART) and *nearest neighbor lookup* for such problems as speech, power security assessment and load forecasting and, in each case, have shown the layered perceptron to perform better in terms of classification or regression accuracy. Both of these competing algorithms can be implemented using parallel processing.

In comparison with nearest neighbor lookup, the layered perceptron was shown to interpolate much more smoothly and with greater accuracy for the problem of power security assessment [1-2].

---

* The majority of the contributions of this paper are not those of the author, but are the result of multitudinous hours of collaboration at the ISDL. The other major contributors are Profs. L.E. Atlas, M.A. El-Sharkawi, J.N. Hwang and Richard Ladner. Drs. Seho Oh and M. Aggoune have also offered significant contributions as have Messrs. Choi and Park. The contributions of these, and others, is better documented in the references.

Here are some accuracy figures contrasting the layered perceptron with CART. Details of the experiments can be found in papers by Atlas *et.al.* [3-4]. Indeed, these papers must be consulted to give significant meaning to the statistics that follow. In power load forecasting [5-6], current and forecasted temperature and current load demand is used to forecast the future power load demand. For this problem, the worst perceptron performance was an error of 1.78%. CART produced an error of 1.68%. For speaker independent vowel classification, the perceptron again had a higher correct classification rate than CART, 47.4% to 38.2%. In the power security assessment problem, the state of a power system is determined to be safe or in jeopardy. Applied to this problem, the perceptron again had a lower error rate - 0.78% to 1.46% [6-8].

In the form of CART used in this experiment, the feature space was initially divided into planes that were perpendicular to the axes. In a higher order form of CART, these planes can be oriented at angles. The higher order form of CART has given preliminary results that are nearly indistinguishable in performance to the layered perceptron. There also exist other high power paradigms, such as *projection pursuit* to which the layered perceptron performance must ultimately be compared.

Query Based Training

More generally, there seem to be fundamental problems of classifiers and regression machines taught by example. These are problems of the problem, and not specific to the classifier used. A dominant difficulty, for example, seems to that of scaling. As is often demonstrated, layered perceptrons work quite well on toy problems such as two bit parity (or the exclusive or). For larger more complex problems, larger more complex layered perceptrons simply will not train. This seems also to be a common problem to other classifiers.

A second problem associated with trained classifiers and regression machines is the diminishing return of information content in randomly generated training data obtained with respect to the data set cardinality. In other words, the more that is learned, the harder it becomes to learn something new. To illustrate, consider the classification problem of learning the location of a point a on the interval $0 < a < 1$. We choose a point at random on the unit interval. If it to the right of a, we assign it a value of one. If is to the left of a, the result is 0. It is clear that, after a number of data points have been generated at random on the unit interval, that a lies somewhere between the rightmost 0 and the left most 1. Call this subinterval C. If we generate a new data point that does not lie in the subinterval C, we have learned nothing new. If the new point lies in the subinterval C, then we revise the subinterval and make it's duration shorter. Doing so, however, decreases the chance that the next data point contains new information. That is, the probability decreases that the new data point lies in the shorter interval. Thus, in this example, the more we learn about the location of the point a, the harder it is to learn. One approach to counteract this phenomenon is with the use of oracles in query based learning [9].

In supervised learning, each feature vector is assigned a classification (or regression) value or values. There is usually a cost associated with this assignment, such as the cost of performing an experiment, computational overhead or simply time. We can envision this process as a presentation to an *oracle* the feature vector. For a cost, the oracle will reveal to us the proper classification or regression value associated with that vector. Note that, if we have deep pockets to pay the oracle, there is no need to for a classifier or regression machine such as the layered perceptron. Any feature vector we desire can be taken to the oracle for proper categorization.

In many cases of interest, we have the freedom to choose the feature vectors that we present to the oracle. Ideally, we would like to present those vectors to the oracle that, in some sense, will result in training data of high information content. The motive is to effectively train the classifier or regression machine with a low training data cost. Query based training is concerned with the manner in which the training vectors that will result in high information data are chosen.

Final Remarks

This paper has posed some important questions in regard to the future success of the layered perceptron artificial neural network and has presented an overview of some of the work being done at the *Interactive Systems Design Laboratory* at the University of Washington to answer these questions.

Acknowledgements

Portions of this work were supported by the *National Science Foundation, Puget Power and Light Company* and the *Washington Technology Center.*

References

1. M.E. Aggoune, L.E. Atlas, D.A. Cohn, M.J. Damborg, M.A. El-Sharkawi and R.J. Marks II, "Artificial neural networks for static system security assessment", **Proc. 1989 IEEE International Symposium on Circuits and Systems,** pp.490-494, 9-11 May 1989, Portland - invited paper.

2. M.A. El-Sharkawi, R.J. Marks II, M.E. Aggoune, D.C. Park, M.J. Damborg and L.E. Atlas, "Dynamic security assessment of power systems using back error propagation artificial neural networks", **Proceedings of the 2nd Annual Symposium on Expert Systems Applications to Power Systems,** pp.366-370, 17-20 July 89, Seattle.

3. L.E. Atlas, R. Cole, Y. Muthusamy, A. Lippman, G. Connor, D.C. Park, M. El-Sharkawi & R.J. Marks II, "A performance comparison of trained multi-layer perceptrons and classification trees", **Proceedings of the IEEE** (in review).

4. L.E.Atlas, J. Conner, D.C. Park, M.A. El-Sharkawi, R.J. Marks II, A. Lippman, R. Cole and Y. Muthusamy, "A performance comparison of trained multi-layer perceptrons and trained classification trees", **Proc. 1989 IEEE International Conference on Systems, Man and Cybernetics,** (Hyatt Regancy, Cambridge, Massachusetts, 14-17 Nov. 1989), pp.915-920.

5. D.C. Park, M. El-Sharkawi, R.J. Marks II, Les E. Atlas & M. Damborg, "Electric load forecasting using an artificial neural network",(submitted).

6. M.J. Damborg, M.A. El-Sharkawi & R.J. Marks II, "Potential applications of artificial neural networks to power system operation", **Proc. 1990 IEEE International Symposium on Circuits and Systems,** 1-3 May, 1989, New Orleans, Louisiana- invited paper.

7. M. Aggoune, M.A. El-Sharkawi, D.C. Park, M.J. Damborg and R.J. Marks II, "Preliminary results on using artificial neural networks for security assessment", **Proceedings of the 1989 Power Industry Computer Applications Conference,** pp.252-258, June 1989, Seattle, WA.

8. M.E. Aggoune, M.A. El-Sharkawi, D.C. Park, M.J. Damborg and R.J. Marks II, "Preliminary results on using artificial neural networks for security assessment", **IEEE Transactions on Power Engineering** (in press). - a reprint of reference 7.

9. M. Aggoune, L.E. Atlas, M. El-Sharkawi, R. Ladner, R.J. Marks II & D.C. Park, "Training connectionist networks with queries and selective sampling", **Neural Information Processing Systems,** (November, 1989), Denver, Colorado.

10. J.N. Hwang, S. Oh, J.J. Choi & R.J. Marks II, "Classification boundaries and gradients of trained multilayer perceptrons", **Proc. 1990 IEEE International Symposium on Circuits and Systems**, 1-3 May, 1990, New Orleans, Louisiana.

11. J.N. Hwang, J.J. Choi, S. Oh, R.J. Marks II, "Query learning based on boundary search and gradient computation of trained multilayer perceptrons", **Proceedings of the International Joint Conference on Neural Networks**, San Diego, June, 1990.

12. J.N. Hwang, C.H. Chan, R.J. Marks II, "Frequency selective surface design based on iterative inversion of neural networks", **Proceedings of the International Joint Conference on Neural Networks**, San Diego, June, 1990.

13. C.M. Lam, D.C. Park, L. Tsang, R.J. Marks II. A. Ishimaru & S. Kitamura, "Determination of particle distribution using a neural network trained with backscatter measurement", **Proc. 1990 IEEE Ap-S International Symposium & URSI Radio Science Meeting**, 7-11 May, 1990, Dallas, Texas.

14. A. Ishimaru, R.J. Marks II, L. Tsang, C.M. Lam, D.C. Park & S. Kitamaru, "Optical sensing of particle size distribution by neural network technique", **Proc. 1990 International Geoscience and Remote Sensing Symposium**, 20-24 May, 1990 , Washington, D.C.

15. A. Ishimaru, R.J. Marks II, L. Tsang, C.M. Lam, D.C. Park and S. Kitamaru, "Particle size distribution using optical sensing and neural networks", submitted to **Optics Letters**.

Biography

Robert J. Marks II joined the faculty of the Department of Electrical Engineering at the University of Washington, Seattle, in December of 1977 where is currently holds the title of Professor. Prof. Marks was awarded the *Outstanding Branch Councilor* award in 1982 by IEEE and, in 1984, was presented with an *IEEE Centennial Medal*. He was Chair of *IEEE Neural Networks Committee* and was the co-founder and first Chair of *the IEEE Circuits & Systems Society Technical Committee on Neural Systems & Applications*. Prof. Marks was also elected the first President of the *IEEE Council on Neural Networks*. He is a Fellow of the *Optical Society of America* and a Senior Member of IEEE. Dr. Marks was also the co-founder and first President of the *Puget Sound Section of the Optical Society of America* and was recently elected that organization's first honorary member. His research interests include signal analysis, detection theory, signal recovery, optical computing and artificial neural processing. Dr. Marks is a co-founder of the *Christian Faculty Fellowship* at the University of Washington. He is a member of Eta Kappa Nu and Sigma Xi.