Title
Network Analysis using Visualization and Singular Value Decomposition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Author
Carbonaro, Mike

Publication Date
1998

Peer reviewed

Network Analysis using Visualization and Singular Value Decomposition

Mike Carbonaro (Mike.Carbonaro@UAlberta.ca)
Department of Educational Psychology; University of Alberta
Edmonton, Alberta T6G-2G5 CANADA

Introduction

The internal response structure for two Parallel Distributed Processing (PDP) networks, used to model a concept attainment task, were compared. These networks' hidden cellular responses were examined using graphical techniques and singular value decomposition. This examination was carried out on both the dynamic and final state of the network produced after network training. One version of the network contained extra output units that constrained the network's learning space and help the network to learn faster and achieve better network generalization (Gallmo & Carlstrom, 1995). It was concluded from the internal analysis of the networks' response structure that the constrained network learned a set of rules which produced greater discrimination among exemplars without any loss to correct categorization.

Internal Network Analysis

McClelland and Jenkins (1991) used diagrams to plot the dynamic learning performance of the network at various developmental stages during the learning of their network's representation. A key part of their visual analysis is the plotting of graphs that show the epoch by epoch performance. Hinton (1986) suggested one could infer certain facts about weight data by visualizing the data in a diagrammatic fashion, often referred to as "Hinton Diagrams." Another analytical technique which appears to be effective in understanding the solution structure of the network is a statistical analysis of the activation patterns for hidden response cells (Hanson & Burr, 1990).

An important factor that can contribute to our interpretation of a network is to create a visual depiction of the internal network representation (Hunka & Carbonaro, 1997). In the case of the work discussed here the interpretation of dynamic learning focused on the changes of the cellular responses at both hidden and output layer. In this context dynamic learning refers to the changes in cellular response values that occurred during the epoch by epoch performance of the network with respect to the input exemplars. For example, given the final state of responses in the network, Figure 1 shows the result of plotting the largest three principal components of W (i.e., $\Gamma^{1/2}W$) based on the largest three eigenvalues (1.74, 1.32, and 0.42) and indicates the dimensions by which the hidden layer discriminates among the input exemplars.

The three axes are labeled for each of the principal components and indicate the range of component values. Most notably, input exemplars 2, 3, and 4 are clustered to the left and define negative concept instances, while exemplars 1 and 5 are clustered to the right and define positive concept instances. Input exemplars 2 and 4 cluster in the bottom right hand corner, both of these contain specific concept attributes. Input exemplar 3 is in the upper left quadrant is also separated out based on a uniquely defined set of attributes with respect to the other input exemplars.

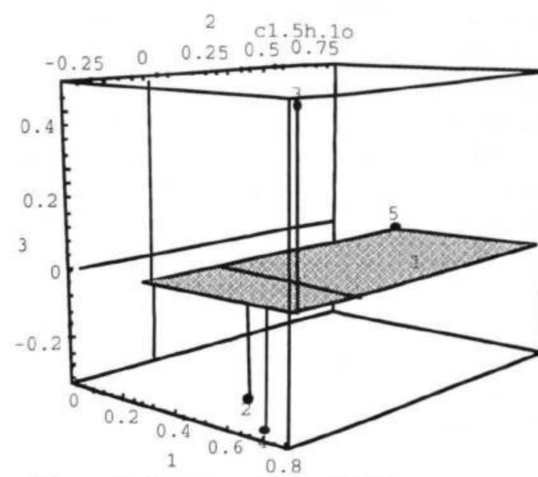

Figure 1: First three PCs of hidden responses

References

Gallmo, O., & Carlstrom, J. (1995). Some experiments using extra output learning to hint multi layer perceptrons. In L. F. Niklasson & M. B. Boden (Ed.), Current trends in connectionism - Proceedings of the 1995 Swedish conference on connectionism, (pp. 179-190). Lawrence Erlbaum.

Hanson, S. J., & Burr, D. J. (1990). What connectionist models learn: Learning and representation in connectionist networks. Behavioral and Brain Sciences, 13, 471-518.

Hinton, G. E. (1986). Learning distributed representations of concepts. In Proceedings of the 8th Annual Meeting of the Cognitive Science Society, (pp. 1-12). Hillsdale, N. J. : Lawrence Erlbaum Associates.

Hunka, S., & Carbonaro, M. (1997). Understanding the back propagated neural network using Mathematica Graphics. Mathematica Education and Research. Spring, pp. 17-21.

McClelland, J. L., & Jenkins, E. (1991). Nature, nurture and connections: Implications of connectionist models for cognitive development. In K. VanLehn (Eds.), Architectures for intelligence: Twenty-second Carnegie symposium on cognition (pp. 41-73). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.