Issues of choosing model complexity and assessing performance and "generalization" (Section 4.3.4) are among the most important open questions. There is some evidence that methods such as cross-validation and AIC are too "local" to fully assess the variability of very flexible methods; therefore some of the assessed benefits of nonlinear methods may be illusory. [On "generalization", Haussler (1992) is a far-reaching extension of the ideas of VCdim to which statisticians, especially David Pollard and Luc Devroye, have contributed; and Anthony and Biggs (1992) is an introductory text on the seminal ideas of Blumer et al., 1989.]

One thing statisticians can contribute to the debate is experience in careful use of sophisticated nonlinear methods. Software is readily available, including in S, and I would encourage statisticians to experiment rather than quote inadequately designed propaganda studies.

To end on a positive note, some very impressive applied statistics is being done using neural networks, and the explosive growth of the subject has opened the eyes of some statisticians (including myself) to the complexity of problems that may be fruitfully attacked by nonlinear methods. I and others have been particularly impressed by some work of my Oxford Engineering Science colleague, Lionel Tarassenko, on analyzing sleep EEG data using both Kohonen nets and radial basis functions to detect structure and anomalous signals (Roberts and Tarassenko, 1993, 1994).

Comment

Robert Tibshirani

Cheng and Titterington's paper is a scholarly overview of the field of neural networks. It should raise the statisticians' awareness of this interesting and important field. One of the authors' objectives was to encourage cross-disciplinary research between neural network researchers and statisticians. Here at the University of Toronto, I have been collaborating informally with Geoffrey Hinton of the Computer Science department, and I think that this collaboration has been fruitful for both of us.

First I would like to make a general point drawing a distinction between statistics and neural networks:

Statisticians tend to work with more interpretable models, since measuring the effects of individual input variables, rather than prediction, is often the purpose of the analysis.

Having said that, there is still much that one field can learn from the other. I will briefly summarize some of the main points:

WHAT THE STATISTICIAN CAN LEARN FROM NEURAL NETWORK RESEARCHERS

1. We should worry less about statistical optimality and more about finding methods that work, especially with large data sets.
2. We should tackle difficult real data problems like some of those addressed by neural network researchers, like character and speech recognition and DNA structure prediction. As John Tukey has said, it is often better to get an approximate solution to a real problem than an exact solution to an oversimplified one.
3. Models with very large numbers of parameters can be useful for prediction, especially for large data sets and problems exhibiting high signal-to-noise ratios.
4. Modelling linear combinations of input variables can be a very effective approach because it provides both feature extraction and dimension reduction.
5. Iterative, nongreedy fitting algorithms (like steepest descent with a learning rate) can help to avoid overfitting in models with large numbers of parameters.
6. We (statisticians) should sell ourselves better.

WHAT THE NEURAL NETWORK RESEARCHER CAN LEARN FROM STATISTICIANS

1. They should worry more about statistical optimality or at least about the statistical properties of methods.
2. They should spend more effort comparing their methods to simpler statistical approaches. They will be surprised how often linear regression performs as well as a multilayered percep-

Robert Tibshirani is Associate Professor, Department of Preventive Medicine and Biostatistics, University of Toronto, 12 Queens Park, Toronto, Ontario M5S 1A8, Canada.