

Neural Networks

MICHAEL I. JORDAN

Massachusetts Institute of Technology

CHRISTOPHER M. BISHOP

Aston University

Neural networks have emerged as a field of study within AI and engineering via the collaborative efforts of engineers, physicists, mathematicians, computer scientists, and neuroscientists. Although the strands of research are many, there is a basic underlying focus on pattern recognition and pattern generation, embedded within an overall focus on network architectures. Many neural network methods can be viewed as generalizations of classical pattern-oriented techniques in statistics and the engineering areas of signal processing, system identification, optimization, and control theory. There are also ties to parallel processing, VLSI design, and numerical analysis.

A neural network is first and foremost a graph, with patterns represented in terms of numerical values attached to the nodes of the graph and transformations between patterns achieved via simple message-passing algorithms. Certain of the nodes in the graph are generally distinguished as being *input* nodes or *output* nodes, and the graph as a whole can be viewed as a representation of a multivariate function linking inputs to outputs. Numerical values (*weights*) are attached to the links of the graph, parameterizing the input/output function and allowing it to be adjusted via a *learning algorithm*.

A broader view of a neural network architecture involves treating the network as a statistical processor, characterized by making particular probabilistic assumptions about data. Patterns appearing on the input nodes or the

output nodes of a network are viewed as samples from probability densities, and a network is viewed as a probabilistic model that assigns probabilities to patterns. The problem of learning the weights of a network is thereby reduced to a problem in statistics—that of finding weight values that look probable in the light of observed data.

The links to statistics have proved important in practical applications of neural networks. Real-world problems are often characterized by complexities such as missing data, mixtures of qualitative and quantitative variables, regimes of qualitatively different functional relationships, and highly nonuniform noise. Neural networks are complex statistical models with the flexibility to address many of these complexities, but as with any flexible statistical model one must take care that the complexity of the network is adjusted appropriately to the problem at hand. A network that is too complex is not only hard to interpret but, by virtue of overfitting the random components of the data, can perform worse on future data than a simpler model. This issue is addressed via statistical techniques such as cross-validation, regularization, and averaging, as well as the use of an increasingly large arsenal of Bayesian methods. Other practical statistical issues that arise include the assignment of degrees of confidence to network outputs (“error bars”), the active choice of data points (“active learning”), and the choice between different network struc-

tures (“model selection”). Progress has been made on all these issues by applying and developing statistical ideas.

The statistical approach also helps in understanding the capabilities and limitations of network models and in extending their range. Neural networks can be viewed as members of the class of statistical models known as “non-parametric,” and the general theory of nonparametric statistics is available to analyze network behavior. It is also of interest to note that many neural network architectures have close cousins in the nonparametric statistics literature; for example, the popular *multilayer perceptron* network is closely related to a statistical model known as “projection pursuit,” and the equally popular *radial basis function* network has close ties to kernel regression and kernel density estimation.

A more thoroughgoing statistical approach, with close ties to “semiparametric” statistical modeling, is also available in which not only the input and output nodes of a network but also the intermediate (“hidden”) nodes are given probabilistic interpretations. The general notion of a *mixture model*, or more generally a *latent variable model*, has proved useful in this regard; the hidden units of a network are viewed as unobserved variables that have a parameterized probabilistic relationship with the observed variables (i.e., the inputs, the outputs, or both). This perspective has clarified the links between neural networks and a variety of graphical probabilistic approaches in other fields; in particular, close links have been forged with hidden Markov models, decision trees, factor analysis models, Markov random fields, and Bayesian belief networks. These links have helped to provide new algorithms for updating the values of nodes and the values of the weights in a network; in particular, EM algorithms and stochastic sampling methods such as Gibbs sampling have been used with success.

Neural networks have found a wide range of applications, the majority of

which are associated with problems in pattern recognition and control theory. Here we give a small selection of examples, focusing on applications in routine use.

The problem of recognizing handwritten characters is a challenging one, that has been widely studied as a prototypical example of pattern recognition. Some of the most successful approaches to this problem are based on neural network techniques and have resulted in several commercial applications. Mass screening of medical images is another area in which neural networks have been widely explored, where they form the basis for one of the leading systems for semi-automatic interpretation of cervical smears. As a third example of pattern recognition we mention the problem of verifying handwritten signatures, based on the dynamics of the signature captured during the signing process, where the leading approach to this problem is again based on neural networks.

An example of a control application based on neural networks involves the real-time adjustment of the plasma boundary shape in a tokamak fusion experiment, which requires several interdependent parameters to be controlled on time scales of a few tens of microseconds. Neural networks have also been applied to the real-time control of the mirror segments in adaptive-optics telescopes, used to cancel distortions due to atmospheric turbulence.

The prospects for neural networks seem excellent, given the increasing sophistication of the underlying theory, the increasing range of applicability of the techniques, and the growing scale of the applications that are being undertaken. The interdisciplinary nature of research in the field seems certain to persist and to bring new vigor into allied fields. Finally, future progress in theoretical neuroscience will provide a continuing impetus for the development and understanding of network models of intelligence.

REFERENCES

- BISHOP, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, New York.
- HAYKIN, S. 1994. *Neural Networks: A Comprehensive Foundation*. Macmillan, New York.
- HERTZ, J., KROGH, A., AND PALMER, R. G. 1991. *Introduction to the Theory of Neural Computation*. Addison-Wesley, Reading, MA.