Chapter V
Visual and Non-Visual Analysis in Archaeology

FROM OBSERVABLE EFFECTS TO UNOBSERVABLE CAUSES

As we have discussed in previous chapters, an artificial neural network is an information-processing system that maps a descriptive feature vector into a class assignment vector. In so doing, a neural network is nothing more than a complex and intrinsically nonlinear statistical classifier. It extracts the *statistical central tendency* of a series of exemplars (the learning set) and thus comes to encode information not just about the specific exemplars, but about the stereotypical feature-set displayed in the training data (Churchland, 1989; Clark, 1989, 1993; Franklin, 1995). That means, it will discover which sets of features are most commonly present in the exemplars, or commonly occurring groupings of features. In this way, semantic features statistically frequent in a set of learning exemplars come to be both highly marked and mutually associated. "Highly marked" means that the connection weights about such common features tend to be quite strong. "Mutually associated" means that co-occurring features are encoded in such a way that the activation of one of them will promote the activation of the other.

As a learning mechanism, a neural network looks as if it explicitly generates and stores prototypes of, for example, the typical stone knife of this period, the typical burial practice in this community, the typical social organization in this period and place. However, there are no such explicit, stored items. What exist are sets of connection weights and synaptic efficacies, respectively. The prototype is not a thing stored at some specific place within the network; it is not an ideal representation of reality waiting to be retrieved by a stimulus. The extraction of the *prototype* arises as an emergent consequence of the proper selection of some characteristic features or input variables.

A prototype as formed within a neural network is by definition "general," in the same sense in which a property is general: it has many instances, and it can represent a wide range of diverse examples. However, this property does not mean that prototypes are universal generalizations. No prototype feature needs to be universal, or even nearly universal, to all examples in the class. Furthermore, prototypes allow us a welcome degree of looseness precluded by the strict logic of universal quantifier: not all *F*s need to be *G*s, but the standard or normal ones are, and the non-stan-

Copyright © 2008, IGI Global, distributing in print or electronic forms without written permission of IGI Global is prohibited.

dard ones must be related by a relevant similarity relationship to these that properly are *G*.

Different neurons represent different "prototypical values" along the continuum, and respond with graded signals reflecting how close the current exemplar is to their preferred value. Note that what is really being stored is the degree to which one neuron, representing a micro-feature of the final concept or prototype, predicts another neuron or micro-feature. Thus, whenever a certain configuration of micro-features is present a certain other set of micro-features is also present (Rumelhart, 1989). This is important, because it means that the system does not fall into the trap of needing to decide which category to put a pattern in before knowing which prototype to average. The acquisition of the different prototypes proceeds without any sort of explicit categorization. If the patterns are sufficiently dissimilar, there is no interference among them at all.

It is clear that a single prototype represents a wide range of quite different possible inputs: it represents the extended family of relevant features that collectively unite the relevant class of stimuli into a single category. Any member of that diverse class of stimuli will activate the entire prototype. In addition, any other input stimulus that is *similar* to the members of that class, in part or completely, will activate a pattern that is fairly close to the prototype. Consequently, a prototype vector activated by any given visual stimulus will exemplify the accumulated interactions with all the possible sources of the same or similar stimuli in proportion to the frequency with which they have been experienced.

The ability to represent both prototypical information and information about specific instances is the basis of the neurocomputing success. We can activate two properties, and discover which outputs are most likely to fit that scenario. The network will initially produce higher activations in the output units which posses any of these properties, with those sharing both properties getting the highest activations. The units for the most widely shared properties also become the most active. Thus the network not only identifies which outputs shared the initial pair of properties, but what their other properties were likely to be, and which among those not possessing the initial pair show the best fit with those who did satisfy the initial pair of properties.

This is an important property of the model, but the importance of this property increases when we realize that the model can average several patterns in the same composite memory trace. Thus, one neural network can be trained to exhibit behavior appropriate to knowledge of a number of distinct prototypes, such as an arrow point, a settlement of a particular kind, or a kind of social organization. Interestingly, if the input is indeterminate between a stone knife and a stone scraper, for instance, the neural network will generate an overall pattern, as if it had an idea not just of knives and scrapers but also on stone tools. We see then that the talent of the system is used to generate a *typical* set of properties associated with some description, even though all the system directly knows about are individuals, none of whom needs to be a perfectly typical instantiation of the description in question.

This way of representing concepts is the consequence of *graduated learning* in a neural network: a new concept emerges as the result of a number of different learning situations or the gradual differentiation of a single concept into two or more related ones. Therefore, as activation spreads from input to output, outputs *grade* according to how well they exemplify the existing training exemplars. Considering that several different prototypes can be stored in the same set of weights, a typical single prototype model may represent instances as sets of attributes (properties or features) with some numeric measure of both the importance of the attribute to that concept (sometimes called its weight) and the extent to which the attribute is present. In this

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/visual-non-visual-analysis-archaeology/6823

Related Content

Enhanced Complex Event Processing Framework for Geriatric Remote Healthcare
V. Vaidehi, Ravi Pathak, Renta Chintala Bhargavi, Kirupa Ganapathy, C. Sweetlin Hemalatha, A. Annis Fathima, P. T. V. Bhuvaneswari, Sibi Chakkaravarthy S.and Xavier Fernando (2018). *Handbook of Research on Investigations in Artificial Life Research and Development (pp. 348-379).*
www.irma-international.org/chapter/enhanced-complex-event-processing-framework-for-geriatric-remote-healthcare/207211

COVID-19 Detection Using Chest X-Ray Images Based on Deep Learning
Sudeshna Sani, Abhijit Bera, Dipra Mitraand Kalyani Maity Das (2022). *International Journal of Software Science and Computational Intelligence (pp. 1-12).*
www.irma-international.org/article/covid-19-detection-using-chest-x-ray-images-based-on-deep-learning/312556

Multi-Objective Particles Swarm Optimization Approaches
Konstantinos E. Parsopoulosand Michael N. Vrahatis (2008). *Multi-Objective Optimization in Computational Intelligence: Theory and Practice (pp. 20-42).*
www.irma-international.org/chapter/multi-objective-particles-swarm-optimization/26950

Supervision of Industrial Processes using Self Organizing Maps
Ignacio Díaz, Abel A. Cuadrado, Alberto B. Diez, Manuel Domínguez, Juan J. Fuertesand Miguel A. Prada (2012). *Intelligent Data Analysis for Real-Life Applications: Theory and Practice (pp. 206-227).*
www.irma-international.org/chapter/supervision-industrial-processes-using-self/67450

CoPBoard: A Catalyst for Distributed Communities of Practice
Gilson Yukio Satoand Jean-Paul A. Barthès (2010). *International Journal of Software Science and Computational Intelligence (pp. 52-71).*
www.irma-international.org/article/copboard-catalyst-distributed-communities-practice/39105