

Decision Support Technique Developed in MATLAB® Improves the Accuracy of Medical Diagnoses

An interdisciplinary research team at the University of Sheffield in England used MATLAB® to develop a data visualization system based on neural networks and pattern recognition techniques. This system is a boon to clinicians, says pathologist Dr. Simon Cross, because it lets them “see immediately where a new case fits into the overall map, and therefore make a more reliable diagnosis.” He adds, “Given the pressure of work that many clinicians are under, a tool such as this can save a considerable amount of time and ensure greater confidence in the results.”

The Challenge

Analyzing and interpreting the huge quantities of available data is a taxing problem for today’s diagnostician. Using traditional spreadsheets and tables, it is difficult to identify the trends, anomalies, and other significant features in the data that are essential to accurate diagnosis.

The University of Sheffield research team, led by Dr. Rob Harrison of the Department of Automatic Control and Systems Engineering and Dr. Simon Cross of the Department of Pathology, set out to create a more effective diagnostic tool. “The human brain is far better equipped to interpret visual images than rows of numbers,” says Cross, a breast cancer specialist. “Our aim was to devise a system that

could convert statistics into images that would be easier for researchers and diagnosticians to analyze.”

Neural networks were ideal for this type of work, explains Harrison: “They can cope with noisy data and, unlike some other systems, they begin by assuming nothing, so they can sometimes identify anomalies that might defeat other systems. They provide a practical, day-to-day tool for reaching the conditional mean without becoming bogged down in theory.”

The Solution

Dr. Harrison, who began using MATLAB in 1986 and has “never coded in anything else since,” recommended MATLAB for this project. Standard neural networks deal with multi-dimensional data, which is difficult for the human brain to assimilate. MATLAB reduces this data to 3-D graphs and plots with overlays that can be readily understood even by inexperienced observers. An additional benefit, says Harrison, is that “MATLAB is effortless to learn and use and lets students get to grips with sophisticated ideas within a very short time.” Dr. Cross, who had never encountered MATLAB before, learned to write short functional MATLAB programs and use MATLAB to manipulate the color graphic output virtually overnight.

The Challenge

To provide clinicians with an effective tool for the analysis and interpretation of large clinical data sets.

The Solution

Use MATLAB to develop a decision support system that presents data as easy-to-read 3-D plots and graphs.

The Results

- A fast, convenient system that refines medical diagnosis
- An effective diagnostic tool for a variety of diseases
- Broad engineering applicability
- Professional acclaim for the new system

“MATLAB contains extremely sophisticated math and engineering functions, and it is easy to add your own. The result is a tool that allows engineers to produce designs in the field and gain a quick, reliable understanding of what they are observing.”

— Dr. Rob Harrison, Department of Automatic Control and Systems Engineering, University of Sheffield

Application Areas

Biomedical
Data analysis

MathWorks Products Used

MATLAB

The technique they developed uses a growing cell structure (GCS) network, an unsupervised self-learning artificial neural network. A graphical user interface developed in MATLAB allows clinicians to quickly start using this method. GCS sorts the data into types (similar cases), which become nodes on a MATLAB graph. No reference is made to the outcome while the data is being sorted. Only after the system has attained its final size is each case linked to its known outcome. Statistical methods are used to place new cases correctly on the graph. The frequencies of cases with each outcome can predict the outcome of any new case placed at that node on the graph.

To test the technique, the researchers selected 692 breast lump specimens, noting the patient's age and ten other pertinent variables. They used either an open biopsy or information from mammograms, combined with the absence of further malignancy, to confirm the final outcome of benign or malignant cancer. The data, which was applied to the neural network, was randomly partitioned into a training set of 462 and a case-test set of 230. Using conventional techniques, they derived a logistic equation from the training set which they then applied to the case-test set.

Next, working in MATLAB, they applied the GCS methodology to the same data. MATLAB produced a network that contained areas of high-certainty positive and negative outcomes, with a narrow area of transition in between. This result compared very favorably with that obtained using the conventional logistic regression technique (96% as opposed to 98%).

“The new technique does not replace human interpretation of data, but it provides a powerful

decision support mechanism to help validate the results,” says Dr Cross. “It is the most useful technique we have had for years.”

The Results

- **A fast, convenient system that refines medical diagnosis.** The GCS technique highlighted specific medical characteristics that were identified quite differently, and inaccurately, by logistic regression. Moreover, GCS took far less time and was much easier for diagnosticians to interpret.
- **An effective diagnostic tool for a variety of diseases.** The system has already been applied to 1000 different cases with three random variables. Tests on lung cancer, heart disease, and two types of bowel disease have all worked well.
- **Broad engineering applicability.** “Neural networks are ideal for engineers, who need reliable results but have limited time to spend checking the fine detail and theory,” explains Harrison. A similar system could be applied to any general fault diagnosis task, such as the identification of anomalous conditions in blast furnace operations or of faults in gas turbine engines—two other projects under way in the Department of Automatic Control & Systems Engineering.
- **Professional acclaim for the new system.** The GCS technique recently won an award for the best presentation in its session at an international conference on neural computing. A paper outlining the technique was published in a recent edition of the prestigious medical journal, *The Lancet* (1999: 354:1518-1521).



3 Apple Hill Drive
Natick, MA 01760-2098 USA
Tel: 508-647-7000
Fax: 508-647-7101
E-mail: info@mathworks.com
www.mathworks.com

To find out more:
www.mathworks.com
www.shef.ac.uk/uni/projects