Neural Nets and Scientific Discovery: A Match Made in AI Heaven

By Ilana Marks

Introduction

The human brain is an invaluable and amazing organ. Electrical and chemical messages bouncing around in the brain lead to actions, speech, thoughts, and all other physical and mental properties that define the human condition. One of the most interesting abilities that the brain affords us is the ability to recognize things, people, scents, and other physical stimuli. When you think about the diversity even within a certain classification of object, it becomes more amazing. For instance, within the classification of tree, there is a wide variety of different types of trees, many of which look drastically different. A pine tree looks very different from an oak tree and yet we realize that they are both trees. We recognize general characteristics of a tree and then are able to expand our definition when other specimens are encountered. We also recognize different states in the life of a tree - for example, we recognize that leaves may change color or fall off. Despite such seasonal anomalies, a tree is clearly recognizable.

In addition to the ability to recognize things, the brain also allows us to predict outcomes and occurrences. When dark clouds appear in the sky, we can predict that rain is imminent (unless, of course, the weather report says it will rain - in that case, the sky will suddenly become clear!) Sports fans try to predict the outcome of a game based on the previous performance of the teams. Given a set of circumstances, it is possible to deduce the most likely result. Typically, the larger the set of circumstances available to make a prediction on, the more likely it is to come up with the correct results.

Recognition and prediction are both highly involved in scientific research. Since science is based on other science it is important to recognize where a result is reminiscent of a previous discovery. Recognizing those connections allows the researcher to expand their understanding by applying information garnered from previous research. A cornerstone of good scientific research is the formulation of an appropriate hypothesis. A hypothesis is nothing more than a prediction about what will occur. Therefore, prediction is a backbone of research. Scientists are constantly striving to increase their knowledge base in order to make more accurate predictions about new experiments.

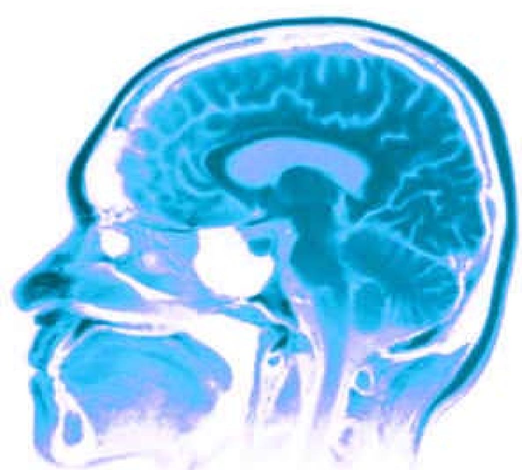

Figure 1: The human brain

Neural Nets and Research

Recognition and prediction may be important to good research, however in this day and age, speed is almost more important. Universities demand that professors publish papers often and the public demands new drugs and technology. The time needed for scientists to learn all about a variety of different fields makes it prohibitive for them to make truly accurate predictions. In addition, a certain type of scientist may not recognize a certain occurrence because it may fit more in the knowledge base of another type of scientist. With the vast increase in scientific discovery over the last several decades, a mountain of knowledge is quickly being created and its growth shows no signs of stopping. The human brain simply cannot deal with so much information. Therefore, it is necessary to employ artificial intelligence techniques in the process of scientific research. Artificial intelligence provides increased speed and "capacity" compared to the human brain, which proves incredibly useful to researchers.

Specifically, neural networks are a very useful tool. Neural networks mimic the behavior of the human brain, so they are often used in applications where systematic thought processes are important. Processing units are interconnected and communicate like biological neurons. Successful and relevant connections are emphasized whereas irrelevant connections are downgraded. This mimics the conditioned learning of biological organisms. This dynamic learning is vital in the research process since science does not remain static.

Cancer Prediction

Cancer is a difficult disease to research. Cancer results from everyday cells that, in effect, "lose their mind." In normal, healthy cells, the process of cell replication is highly regulated. Many "checkpoints" are in place to make sure that all steps are done properly. These "checkpoints" are usually molecules that are created or degraded which signal the cell to either hold in the stage where it is or to move on to the next stage of cell division. For example, one such molecule believed to be involved in several different cancers is called p53. This molecule is often absent or mutated in extracts from tumor cells, indicating that perhaps p53 is a molecule that signals the cell to stop dividing. If p53 is not produced, then cells will divide too rapidly and cancer can result.

Since no two cancer physiologies are exactly alike, it is notoriously difficult to make a prediction about whether the patient is likely to survive. Even with two cases of the same type of cancer, one person may survive for many years while another will die within months. A myriad of different physical factors contribute to the outcome. This is one reason why cancer is such a frightening disease - no one can tell you whether your particular complement of genes will put you at an advantage or disadvantage.

Researchers at the National Cancer Institute (NCI) are working on creating a model that will help predict the prognosis of cancer patients. Particularly, they are working with a type of cancer called Neuroblastoma. Neuroblastoma is a childhood cancer. It usually begins with cells of the adrenal gland and spreads, creating tumors in the neck, chest, abdomen or pelvis. Using DNA microarray analysis (see PC AI Volume 18, Issue 2 - "Microarrays and Artificial Intelligence") the researchers studied the gene expression profiles of cancer patients. The microarrays consisted of about 25,000 genes and the analysis was repeated for 49 patients. In order to connect the microarray results with certain outcomes, the 49 patients were chosen because their outcomes were known. Some of the patients survived for more than three years without any cancer-related issues. Others died due to the disease. Feeding this information into an artificial neural network, the researchers found they