Product Detail

Traditional approaches to analyzing data usually adopt numerical techniques that apply a variety of statistical principles. These approaches require substantial expertise and judgment especially when making assumptions. The inherent nature of making assumptions means you are discounting possible results before you begin. RapAnalyst makes no assumptions about the nature of the interrelationships present in data and therefore is more accurate in successfully identifying them.

RapAnalyst allows you to visualize and work with high dimensional data (i.e. many variables) within the two-dimensional Knowledge Filter™. A Knowledge Filter is an optimized representation of the input data that is constructed during the RapAnalyst training process. This training process, which uses unsupervised neural networks and genetic algorithms, arranges a representation of the input data in the Knowledge Filter based on similarity.

The Knowledge Filter is comprised of a large number of connected hexagons called nodes. All relevant variables considered, two nodes close together are more similar than two nodes further apart. The same Knowledge Filter can be viewed for any variable in the data and are called attribute window views. When several different attribute windows are opened simultaneously the investigative power of RapAnalyst truly emerges.

The attribute windows are a color depiction of complex multi dimensional data in two dimensions – a bird's eye view of the data. As each attribute is displayed in its own window, you can easily identify the dynamics and interrelationships within the data. This depiction can also quickly provide insight into how and why certain events occur. Once your data is represented in the Knowledge Filter, RapAnalyst allows you to visually and quantifiably investigate, view grouping statistics, run predictive analysis and test scenarios with your data.

RapAnalyst brings to every decision maker the power of high prediction accuracy and illustrates the interrelationships between all attributes within your data. The Knowledge Filter can also predict missing values interactively for single cases, predict any number of cases using external files in a batch mode and will automatically predict values for missing data when presented with incomplete data, all easily accessible by the non-technical user in real-time.

The Cluster functions of RapAnalyst logically divide your data into an assessment of the "best" grouping as guided by a very simple approach: make everything within the same group as similar as possible, while making everything between different groups as dissimilar as possible. Clustering has a natural application for customer segmentation and any other requirement to identify the most optimized groups present in your data. A Knowledge Filter can be used to quickly decide in which group each record in new, emerging data belongs.

Statistic views provide the ability to quantitatively substantiate features that are seen in the Knowledge Filter. Statistics are available for every attribute of all nodes, a single node, neighborhood of nodes, any combination of nodes or any cluster in the Knowledge Filter.

RapAnalyst utilizes delimited text files or Microsoft Excel worksheets to build each Knowledge filter. The filters are created by applying sophisticated machine learning techniques based on a proprietary unsupervised neural network algorithm to high dimensional complex data. RapAnalyst's learning is fast, assumption free, tolerant of imperfect data and does not need formulas or rules. Ease of use was the key driver behind the development of RapAnalyst empowering the decision maker to understand, effectively interpret and communicate the results of data analysis to a wide audience, without reliance on specialists. During the learning process RapAnalyst takes a total approach to representing the data considering all attributes equally in order to develop an accurate understanding of the interrelationships between all attributes within the data, this is in contrast to most traditional analytical approaches which target specific

attributes that are assumed relevant. Data records are arranged in the model based on their similarity, all things considered.

Neural Network Background and Overview

Neural networks are the result of research attempts to model biological brain functions. They have the ability to learn relationships between data, without a prior understanding of an underlying problem area. Neural research started in the 1950s, and has intensified ever since. The technology around neural networks is divided into supervised and unsupervised neural networks. The fundamental differences between the two types are as follows:

> *Supervised* – These neural networks require the user to define a specified outcome for any learning to take place. This network will learn the relationships between input and output attributes.
>
> *Unsupervised* – These neural networks make no assumptions about input and output attributes. Each data record is regarded as a single unit. The neural network clusters data records based on similarity. The network develops relationships between the attributes based on the clustering.

The neural networks deployed by RapAnalyst fall into the category of unsupervised neural networking. Traditionally, neural networks represented a “black art” practiced only by scientists and researchers, and were not for the average person. Recently, the commercial possibilities became apparent, as research based work was applied to real-world applications.

One of the most widely used applications of neural networks in the business world is data mining. Data mining is the process of extracting knowledge from information. It is often policy in companies to collect vast amounts of data about every aspect of its business.

Useful information, such as customer surveys, churn rates, expenses, etc. are available, but unfortunately very little is done with this data because of its complex nature. If business data cannot be understood it has little value, and is very often ignored. A lack of useful, understandable data can lead to decisions being made on vague assumptions alone. The risks associated with this course of action can be disastrous.

Neural networks alleviate this burden by providing decision support that is based on solid and proven research. With RapAnalyst the development of an application that encapsulates neural networks makes this technology more accessible to non-scientific users.

Advantages over other techniques and systems

The type of neural network used by RapAnalyst does not require any domain specific information to create a possible solution to a problem. This no-assumption model allows for processing information without possible inaccurate operator guidelines. The biggest advantage of no-assumption approaches is that neural networks search for any interdependency among data components. Certain possible relationships may not be obvious even to a seasoned data analyst.

Neural networks are however not hindered by experience, and will be more likely to find relationships that weren’t apparent to their human counterparts. Neural networks also excel in finding non-linear relationships between different attributes. These relationships are by nature complex, and humans may easily miss them without the use of specific tools.

It is important to note that RapAnalyst does have a number of related applications within the scientific community. A scientific user with powerful statistical knowledge and background will, with time, be able to build simulators that will satisfy particular problem areas. A major shortfall with this approach is that is up to the user to design the statistical approach employed for the particular problem. This methodology is based on an experimental approach to a particular data problem. The end result has meant that few of these potentially powerful applications have proven viable outside of the scientific environment.

The RapAnalyst Neural Network

The type of neural network used by RapAnalyst was inspired by how the various human sensory impressions are neurologically mapped onto the brain cortex. In the brain, spatial relations among stimuli correspond to spatial relations among neurons. In a trained unsupervised neural network, similar data

patterns are represented in topologically similar positions on an output plane. This clustering process is not the result of matching inputs to a predetermined classification structure. The trained unsupervised neural network is both a clustering of similar elements and an ordering of the clusters within the observation space. Any clusters that are topologically close to each other will exhibit similar characteristics.

Traditional clustering algorithms sort databases on individual elements of data vectors. Unsupervised networks cluster data by considering complete data vectors. In terms of reward program analysis, a single individual in the population will be considered as a whole instead of using particular attributes. This allows the algorithms to maintain relationships between individual vector elements. These relationships can then also be effectively visualized by constructing a quantized vector field; a Knowledge Filter.

Knowledge Filters are constructed by continually testing and refining data records and understanding the association, similarities and interrelationships between the attributes within the data. No assumptions are made about existing structures in a data, or any statistical relationship. The refinement process leads to a sorted state where similar data vectors will be associated with positions on the knowledge filter that occur in close topological proximity. Sufficient repetition will lead to highly sorted state, which is an excellent approximation of the original data. These networks also have the added advantage that they normally fail to create a good representation if the provided data does not represent real, organized process. This helps identify a problem area within your data capturing method. An algorithm that extracts 'knowledge' from random data would have little usefulness.

Details of the RapAnalyst approach

The construction of a knowledge filter is based on the interaction between two layers. The first layer is an input layer structured around a multidimensional vector representing all the parameters related to the problem under consideration. The second layer is called the output layer, in RapAnalyst this is the Knowledge Filter. The Knowledge Filter is a two dimensional grid. Each intersection defines a node which is also called a neuron. The array has the neurons fully connected to the input layer. This means that there is a link from every element in the input vector to every neuron in the knowledge filter layer.

The knowledge filter performs a projection between the input space and the output space so that the relationships that exist between the input data are reflected on the output layer.

By virtue of this learning algorithm, the knowledge filter performs a non-linear regression. The unsupervised learning process is an iterative refinement of the position associated with each data record. At the end of the learning process, the output grid can be visualized by coloring individual data record attributes for each of the output nodes.

This visualization will highlight the organization of data: Similar attribute values will be visualized using similar colors, thus illustrating the clustering tendency.

The main properties of RapAnalyst neural networks are:

- The mapping of input data to the output plane represents the full set of data in an ordered form. Similarities in data samples will be represented as geometric relationships on the output plane.
- The structures in the data can automatically be visualized on the output plane whereby the degree of clustering may be represented by color changes.
- The natural distribution of the data samples enables the output plane to be used as a natural framework, on which the individual statistical indicators can be visualized as color levels making complex relationships clearly visible on a 2-dimensional plane.
- The process does not require complete data therefore the problem of missing values is resolved.

RapAnalyst's visualization model simplifies complex neurological data relationships into two-dimensional grids that can be easily interpreted. Each attribute from the original data is visualized on its own grid – making it easy to evaluate and identify attribute relationships. This simple interactive process replaces the traditional approach that would have required a data analysis expert to browse through endless tables of numbers.

RapAnalyst can be used in any industry to process seemingly incomprehensible data into sensible business knowledge. The no-assumption approach makes it ideal to extract relationships to be used in decision support, significantly decreasing common risks associated with business decisions.

Particular benefits of RapAnalyst

The real power of RapAnalyst is its ability to merge power with simplicity. This is highly relevant when you consider all the areas where this neural network technology would be applicable. The reason these areas have not had success is that the technology has remained inaccessible due to its complexity. Therefore the most compelling feature of RapAnalyst would not be in its technical power but its ability to take the neural network technology to the actual users of the business process.

Varying degrees of success in the modeling process are experienced when different data models are used. This means that the organization would ideally need an application that allows multiple models and scenarios to be generated quickly and effectively. This is a requirement that RapAnalyst lends itself to well.

RapAnalyst is able to integrate into the real-time decision models of the organization. The application has the ability to “question” data with real-time output results. This will become an ever-increasing requirement as the technology becomes more established in the commercial environment.

Artificial intelligence as a part of a business solution is continuing to grow. The momentum at which this is occurring is also increasing. Management is continually looking for new ways to derive more value. This means that the workplace is undergoing an innovative drive, which sees them re-evaluating so-called "accepted business processes".

About Raptor International

Raptor International is one of the world's leading SOM software developers. With offices in the United States, Australia and South Africa, Raptor helps organizations across the globe effectively conduct rapid data analysis of large, multivariate data sets. Raptor also designs analytical applications that enable companies to better predict future business trends and make critical business decisions based on their own data. Raptor helps companies transform their complex data into actionable information.