10

Evaluating Medical Decision Making Heuristics and Other Business Heuristics with Neural Networks

Steven Walczak

The Business School, University of Colorado at Denver and Health Sciences Center, Campus Box 165, PO Box 173364, Denver, CO 80217 USA
Steven.Walczak@cudenver.edu.

Summary. Heuristics are an efficient means for solving complex and also partial information business problems. Unfortunately, the development of new heuristics and the evaluation of existing heuristics is a labor intensive process. Neural networks provide a fast and reliable method for evaluation of new heuristics against existing heuristics and the optimization of new heuristics when no prior heuristic exists. This chapter describes a methodology for utilizing neural networks as a heuristic evaluation mechanism and discusses how existing research has been utilized (possibly unintentionally) in the development or evaluation of new heuristics.

10.1 Introduction

Businesses face difficult problems on a daily basis. Historically, heuristic methods have been used to solve complex and nonlinear business problem types, such as frequently occur in medicine, finance, and other business areas. While heuristics provide an agile means to quickly evaluate complex independent variables to make a decision for a specific business problem, heuristics by their nature may not always provide an optimal solution (Pearl 1984).

For the purposes of this chapter, a heuristic is defined as any "rule of thumb", that is a decision making model, that enables rapid decision making in domains where complete information regarding the decision making problem is either very complex and thus difficult to evaluate or may be missing. A heuristic is therefore a decision making rule of some type. Elements of heuristics are the information, variables, that are available and are evaluated to form the heuristic decision. Hence, while the variables themselves are not heuristics, they are a necessary criteria for employing a heuristic decision making method and will be treated as equivalent to the heuristic method itself in this chapter.

An interesting question then is how businesses and services may evaluate current and new heuristics or heuristic variables to determine their

S. Walczak: *Evaluating Medical Decision Making Heuristics and Other Business Heuristics with Neural Networks*, Studies in Computational Intelligence (SCI) **97**, 259–287 (2008)
www.springerlink.com

ability to select optimal or near optimal solutions. Of course, proving that a heuristic is optimal is a much harder problem, but evaluation techniques are meant to build confidence in the reliability of the heuristic being evaluated and not to prove optimality. Traditional statistics provide a number of tools for performing such evaluations (e.g., mean error rate, regression modeling, and discriminant analysis), but these evaluation methods suffer from rigid a priori requirements on population distributions and error distributions (Daniel 1999), which frequently are unknown for traditional medical and other business problems.

When a priori error, population, and variable distributions are unknown, nonparametric methods need to be applied (McLachlan 1992). Neural networks are a nonparametric methodology, which means that they model underlying equations without any requirement for a priori knowledge of population or variable distributions.

Although neural networks are often described as a "black box", indicating that information concerning the effect of individual variables on predicted outcomes is difficult to ascertain, they have several advantages that make them an ideal choice in evaluating (as well as developing) decision making models in medicine and business. These advantages are:

- Extremely fast, once trained
- Tolerant of moderate amounts of noise in data
- Prediction and classification models are learned from the data dynamically, producing optimal or near-optimal models

While the idea of a heuristic model is to use a smaller and possibly less expensive set of variables to produce accurate decisions, another advantage of neural networks is in their ability to handle not only noisy data, but also missing data for the heuristic model's variables (Markey et al. 2006).

The possible difficulty of determining the effect of individual decision variables for heuristics presents a problem, this problem may be overcome in several ways as will be discussed in this chapter. The focus of this chapter is on the utilization of neural networks as a tool for evaluating either new heuristic decision models or more frequently evaluating the utility of new or different combinations of variables to either create a new heuristic model or improve upon the performance of an existing heuristic decision model.

The remainder of the chapter is organized as follows, First, factors that impact the performance of neural networks and values for neural network model building are examined. Next, a process for utilizing neural networks to evaluate new heuristic decision making models, including new combinations of variables that serve as the input to the heuristic method, is presented and discussed. Finally, historic examples of the use of neural networks to discover new heuristics or invalidate existing heuristics are presented for both medical and business domains.

10.2 Neural Network Parameters

The advantages of neural networks just mentioned appear to make them an ideal choice for both developing heuristic decision models as well as evaluating existing and new heuristic decision models.

10.2.1 Training Time and "Black Box" Nature

A disadvantage noted in the literature which is no longer true is that even though neural networks are extremely fast, even instantaneous, once trained, the training time can require a large time investment. Current neural network shell programs that assist researchers and modelers in developing neural network solutions typically require only a few minutes to train a moderate size network (less than 100 processing nodes and 2,500 connections). As such the training time cost is negligible for most heuristic decision making problems, since the utilization of heuristics simplifies the decision space. Of course as the size of the neural network architecture grows, the training time requirement will increase correspondingly.

The other significant disadvantage of utilizing neural networks for evaluating the applicability and effectiveness of business heuristics is the "black box" nature of neural networks. This is a disadvantage in evaluating heuristics, since the effect of the variables belonging to the heuristic technique may be difficult to determine. Section 2.3 will discuss a means for overcoming this handicap, but right now the ability to produce an algorithm where independent variables are defined is examined. The "black-box" nature of neural networks is being challenged and various techniques are being developed to mine the variable relationships that drive neural network performance (Zhang 2007) and produce if-then decision heuristics from neural network connections (Ballini and Gomide 2002/2003).

The small neural network shown in Fig. 10.1, utilizes two independent input variables, x_1 and x_2 and produces either a prediction or classification value y. An aggregation algorithm specified by the neural network developer is

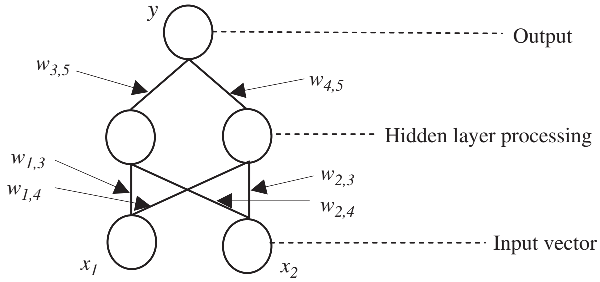

Fig. 10.1. Sample neural network

used to combine input values into a processing node, $f(x_1,\ x_2)$, and a transfer function, $g(X)$, applies an activation function for transferring the nodes value to the next layer in the neural network. Thus:

$$\bullet\ y = f_5(g_3(f_3(x_1, x_2), g_4(f_4(x_1, x_2)), \tag{10.1}$$

where the subscripts indicate the aggregation or transfer function of a particular processing or output node in the neural network. If a simple weighted summation is used for all $f(x_1,\ x_2)$, and a sigmoid is used for all $g(X)$, then (1) becomes:

$$\bullet\ y = \left(\sum_{i=3}^{4} w_{i,5} \frac{1}{1 + e^{-\sum_{m=1}^{2} w_{m,j} x_m}} \right) + \varepsilon. \tag{10.2}$$

The range of i in (2) is the number of hidden processing nodes and the range of m is the number of input nodes. This equation is for a single hidden layer. Additional hidden layers would mean aggregating the weighted sum of the sigmoid function applied to separate copies of (2) for each additional layer processing node. As can be seen, while it is possible to generate an equation and determine variable impact, the size of the equation quickly becomes problematic for understanding the relevance of an individual variable (or collection of variables) as additional processing nodes and hidden layers are added.

An alternate means to evaluate the impact of input variables on the model's output is to examine the values for each connecting weight, $w_{i,j}$. The backpropagation algorithm learns to model domain problems by automatically adjusting these weights to approximate an optimal model, hence neural networks may themselves be seen as heuristic models. Every connection weight in the model must be examined to try and determine the overall influence of specific variables, with the absolute value of the weight serving as an indicator. This technique, like the algorithm generation technique just described is very difficult to accomplish and becomes more problematic as the size of the neural network's hidden node architecture increases.

One final technique and one that is commonly performed by existing neural network shell tools and that is easier to perform is the leave-one-out methodology. This technique emulates the step-wise portion of step-wise regression modeling. Leave-one-out, as the name describes, involves dropping or alternately adding a single variable and comparing the performance of the two corresponding neural network models that differ by only the single variable. If a significant increase or decrease in performance is noted, then the effect of the missing or added variable may be estimated. A drawback of this process is that other interaction effects are likely to be occurring in the neural network and so the final difference in model results is not solely due to the single left out variable. This technique will form a part of the neural network heuristic evaluation process described in the next methodology section.

10.2.2 Other Factors Affecting Neural Network Efficacy for Heuristic Evaluation

Two other factors affecting neural network model development that are debated in the research community, which may turn out to be advantages for using a neural network based nonparametric modeling approach are: training population size and input vector size.

Typically, especially in financial and medical domains, researchers prefer to have larger sample populations for developing and validating decision models (Zahedi 1996; Zhang and Hu 1998). Larger data populations improve statistical significance of the resulting model by increasing the probability that unusual cases are included in the sample population (Schürmann 1996). Recent research has indicated that, at least for the populations studied, much smaller training populations still enable neural network models to achieve very high prediction and classification accuracy rates (Abdel-Aal 2004; Shin et al. 2005; Walczak 2001). Based on these reported research results, it may be possible to utilize neural networks for heuristic model evaluation with much smaller sample (training) sizes than are required for validation utilizing traditional statistical methods.

Even with a very small population of historic data, which is needed for the supervised learning methods, reasonable approximations of the performance of neural network heuristic models may be accurately evaluated utilizing several techniques. Bootstrapping and the jackknifing refinement of bootstrapping (Efron 1982) enable the utilization of members in the historic population set to be utilized as both training and evaluation cases through utilizing a leave one out methodology (Tourassi and Floyd 1997). Similarly, cross-fold (commonly called N-fold) validation is a similar method to jackknifing but which utilizes larger sets of randomly selected evaluation cases from the population and is commonly used in algorithmic validation experiments (Lim et al. 2000).

The other design issue debated among neural network researchers is the relative advantage or disadvantage of learning with respect to input vector size. Early researchers claimed that an advantage of the fact that neural networks learned connection weights was that large quantities of independent variables could be included in a neural network model's input vector and that the neural network learning algorithm would deselect noncontributing variables out of the model (see e.g., Hertz et al. 1991). This could be advantageous, if it works, in that variable additions to heuristics could be tested by adding them in and determining if the neural network has deselected any of the new heuristic variables. Alternatively, if the neural network keeps a heuristic variable over more traditional variable (through the process of deselection) this could serve as validation of the new heuristic.

However, the other side of the input vector size argument claims that each input vector variable impacts the resulting classification or prediction made by the neural network and hence the selection of input variables to the neural network model is a critical decision factor and determines the optimality of the corresponding neural network model (Güler and Übeyli 2005; Nath et al. 1997; Piramuthu et al. 1994; Smith 1993; Soulié 1994; Tahai et al. 1998; Weigand and Zimmermann 1995). In fact, Pakath and Zaveri (1995) claim that input variable selection sensitivity is not only a factor for neural networks, but for other artificial intelligence modeling paradigms as well. This too can serve in the goal of heuristic evaluation, but does remove the simplicity of automation.

The author of this chapter follows the later design criteria, which typically involves the use of domain experts to determine if potential variables could contribute to a heuristic solution to the problem being modeled. One consideration is that highly correlated variables must be removed, so that only the influence from a single variable (or set) affects the neural network model's outcome (Smith 1993), which involves some additional up front analysis of the variables being used.

Additional design considerations must be determined by the researcher desiring to evaluate decision heuristics, such as the type of training algorithm used and the architecture of the neural network model, but will not be discussed at length here as they have already been discussed in detail in the literature (for tutorials on neural network implementation and design issues see (Jain et al. 1996; Rodvold et al. 2001; Zhang 2007)). The remainder of this chapter will focus on supervised learning methods for neural network training, including the very popular backpropagation algorithm, though the techniques described are applicable to unsupervised learning methodologies as well.

10.3 Methodology for Evaluating Business/Medical Heuristics

The advantages of a learning system that dynamically determines classification or prediction equations, is noise tolerant, and may be able to be used with relatively small amounts of data provides an interesting potential for evaluating the potential of business heuristics. As noted in Sect. 2.3 though, determining the contribution of specific input vector variables is problematic and researchers are advised to limit the quantity of input vector variables to be able to more critically evaluate the effect of the heuristic model's variables.

Since input vector variable selection is viewed as a critical step in the development of neural network models, the heuristic variables may be evaluated by adding them to a current set of decision variables and possibly deleting correlated variables (which may be more difficult or more costly to acquire) and evaluating if any classification or prediction performance improvement has occurred, similar to the leave-one-out method described earlier. Typically, this methodology assumes that an existing decision model already exists. If no existing method is available for comparison, then it simply becomes a case of evaluating a new model, regardless of the heuristic nature, and confidence values for the new models output should suffice.

When a competing model is already in use, then two neural network models are required. Since typically a new heuristic decision model will utilize a different number of input variables, the architecture for each model must be optimized independently, following traditional neural network development protocol. Optimizing the number of hidden nodes in each layer and the number of layers is critical for comparing optimal results for each method. The methodology for comparing a new heuristic model is displayed in flowchart format in Fig. 10.2.

The two critical parts of the flowchart in Fig. 10.2 are in the two decision diamonds. Selecting an appropriate statistical method for comparing the output of the two models is important since classification models and forecasting models require different analysis methods and selecting an inappropriate

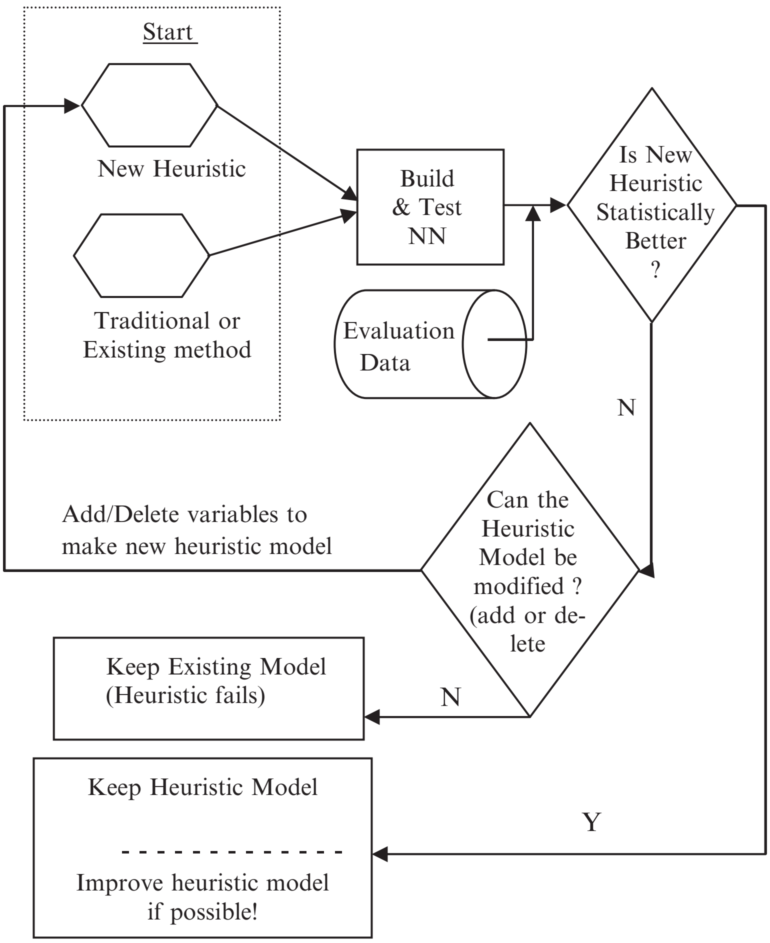

Fig. 10.2. Flowchart of new heuristic comparison methodology

statistic for comparison could significantly alter the evaluation results. In heuristic medical models that involve patients, the sensitivity and specificity of the model's output is the gold standard for comparison, however trade-offs in sensitivity and specificity may be possible. Furthermore, for medical domains, it has been recommend that all neural network models be evaluated using clinical trials or randomized control trials of the neural network's model of the heuristic (Lisboa and Taktak 2006) to demonstrate real world performance of the heuristic model.

The other important aspect of heuristic development is the ability to modify the heuristic through either the inclusion of new decision variables or the exclusion of current variables in the heuristic model. Since the goal of a heuristic is to simplify the decision making task in a complex decision environment, then improving the heuristic's performance through the addition or deletion of variables makes logical sense. In fact, even when a new heuristic model is validated using the method just described, improvements to the model and it's classification or forecasting performance should still be sought via further modification of the input variable set.

Additional modification of an accepted new heuristic follows a path similar to that of comparing a new and existing heuristic. Additional variables may be added or removed to the accepted new heuristic model to develop an improved new heuristic model. This process may be repeated iteratively until all desired combinations of independent variables have been evaluated and the best performing heuristic model is kept (Swanson and White 1995, 1997). The addition or removal of variables in a step-wise fashion aides in determining the effect of individual variables or sets of variables on the heuristic decision model.

For example, a heuristic to evaluate the morbidity of a patient by utilizing four variables for heart rate, breath rate, perspiration, and skin temperature (Tahai et al. 1998) may be able to determine if the patient is in need of CPR (cardio-pulmonary resuscitation) with 93% accuracy, which appears to be a fairly good heuristic classification model. However, removing the last two variables: perspiration and temperature, increases the heuristic model's accuracy to almost 100%. Further reductions in the input variable vector set reduce the overall accuracy of the neural network model to approximately 50%. Addition of a systolic blood pressure variable to heart rate and breath rate variables would also serve to reduce the model's accuracy since systolic blood pressure is correlated with heart rate. Thus, even though an acceptable heuristic model is developed initially, further evaluation should still continue on the variable set used in the heuristic model to possibly improve performance.

Furthermore, neural network heuristic decision model refinement through the deletion of correlated or other unproductive variables assists in reducing the overall data collection costs associated with the neural network (Bansal et al. 1993). The reduction of data collection costs will be especially important in medical domains, if the remaining variables are available through less costly and especially less invasive medical test.

Next we will examine several applications of neural networks in evaluating heuristics and development of new heuristic methods.

10.4 Neural Network Evaluation of Medical Heuristics

Heuristics are frequently used in medical domains where rapid assessment and treatment reduces medical costs and may improve patient outcomes. Sophisticated tests exist that are capable of accurately gathering information about a patient's condition, but these tests are costly and take time to perform. The usage of neural networks to evaluate hypotheses (heuristics) of nonlinear components in exploratory medical data analysis has been previously suggested (Lisboa 2002). However, hypothesis testing with neural networks has thus far only been performed in an ad hoc manner without a formal methodology for evaluating new heuristics.

This section examines several previous research efforts that have been able to validate new decision heuristics or occasionally invalidate existing heuristics in medical resource allocation and diagnostic medical decision problems. The effective management of resources and patient information is critical for managing costs and improving the quality of patient care (Buchman et al. 1994). It should be noted that the original purpose for most of the reported research in this section was to create decision support tools for use in hospitals to manage resource allocation problems or improve diagnosis of specific medical and trauma conditions. The creation of new heuristic methods or the invalidation of existing heuristics occurred as a side effect of the original research, but serves to illustrate the usage of neural networks to evaluate new diagnostic or resource allocation decision heuristics.

Table 10.1 provides a summary of the neural network research that has indicated the development of a new heuristic or the invalidation of an existing heuristic method in a medical domain. The reported prior research in Table 10.1 utilized a variety of supervised learning algorithms and hidden layer architectures, thus demonstrating the robustness of supervised learning neural network training methods for heuristic decision model development and evaluation in medical domains. The training methods specifically include: one and two hidden layer backpropagation or multilayer perceptron, radial basis function, fuzzy ARTMAP, probabilistic neural networks, soft max discriminant analysis, self-organizing maps (an unsupervised learning method), and others.

The research listed in Table 10.1 are cases were a neural network model produced a new heuristic or modified the variable elements of an existing heuristic to create a better performing heuristic decision model. The research listed in Table 10.1 is not meant to be exhaustive (since well over 100 medical domain neural network journal research articles have been published in the last 2 years alone), but rather representative of much of the current medical domain research being done with neural networks.

Table 10.1. Samples of neural networks (NNs) discovering medical heuristics

Medical domain	Heuristic results	Citation
Blood/transfusion resource allocation	MSBOS heuristic previously used for determining blood needs has weaknesses, NNs are better at predicting required blood for transfusions	(Walczak and Scharf 2000)
	Information available upon arrival at an ER could be used to predict the transfusion needs of adult trauma patients	(Walczak 2005)
Brain/epilepsy/ head injury	NN using Doppler velocity variables accurately predicts head injuries	(Erol et al. 2005)
	Uses Lyapunov exponents with EEG to classify epilepsy	(Güler et al. 2005b)
	NNs used to show that EEG rhythmicity may be used to classify seizures as epileptic or nonepileptic	(Nowack et al. 2002; Walczak and Nowack 2001)
	NN analyzes very large amounts of time domain and frequency domain EEG data	(Srinivasan et al. 2005)
Breast cancer	NN utilizes new variables (urokinase-type plasminogen activator (uPA), and plasminogen activator inhibitor-type 1 (PAI-1)) combined with gene expression signatures for determining chemotherapy	(Harbeck et al. 2007)
Heart/heart disease/ circulation	NN uses a large number of available variables, but improves over physician and other IT methods	(Baxt 1991, 1995; Baxt and Skora 1996)
	Establishes a lower cutoff value for the troponin l variable	(Eggers et al. 2007)
	Uses signal/noise ratio to improve NN performance	(Güler and Übeyli 2005)
	Recommends inclusion of ECG data with standard perfusion scans for a more accurate model	(Gjertsson et al. 2006)
	NN used noninvasive variables only to diagnose heart disease	

Table 10.1. (continued)

Medical domain	Heuristic results	Citation
	Evaluated the use of genetic variant variables for predicting venous thrombosis	(Mobley et al. 2005) (Penco et al. 2005)
	Use of $FSPO_2$ and fetal heart rate variability variables with CTG tracings improves diagnosis of fetal hypoxia	(Salamalekis et al. 2006)
Injury severity/ length of stay	Using presentation data only, NN accurately predicted morbidity and length of stay for burn patients	(Frye et al. 1996)
	Using only data from the ER, NN predicts morbidity	(Izenberg et al. 1997)
	Accurate prediction of pediatric acuity of care with information available in the first 10 min of arrival at the ER	(Walczak and Scorpio 2000)
	NN uses noninvasive variables (spectrometry) to indicate burn severity	(Yeong et al. 2005)
Lung	NN uses of new technique, minimal-polyline approximation, to detect emphysema compared to curvature-based features, from chest radiographs	(Coppini et al. 2007)
	NN uses new technology to diagnose lung sounds	
	NN demonstrated that inclusion of a reactive glucose variable would improve the existing heuristic of using the d-dimer value in isolation to predict pulmonary embolism	(Güler et al. 2005a) (Walczak et al. 2006)
Pancreatitis (acute)	NN with less costly variables is able to outperform an abbreviated version of the RANSON score commonly used for predicting acute pancreatitis patient outcomes	(Pofahl et al. 1998; Walczak et al. 2003)
Prostate	NNs enable the use of lower PSA level to reduce false positives and reduce unnecessary biopsies	(Reckwitz et al. 1999)
	NN uses percent free prostate specific antigen variable to improve diagnosis of prostate cancer	(Stephan et al. 2002)

In order to examine the representative neural network research presented in Table 10.1, the cited research will be classified into three categories of new heuristic development:

- Development of a new heuristic when no competing heuristic exists
- Development of a new heuristic that utilizes significantly different variable elements from an existing heuristic
- Development of a new heuristic which is mostly an improvement of an existing heuristic, primarily through modification of the variable elements in the existing heuristic

Much of the research listed in Table 10.1 follows the evaluation methodology shown in Fig. 10.2 by creating the new heuristic model and determining it's performance through a neural network implementation. These neural network models are then compared statistically against the existing heuristic model to determine if the newer model improves overall performance. The neural network performance is evaluated against the previous models which may include statistical models, such as multiple regression, or expert physician performance on the same domain problem. Recall that in medical diagnostic domains, performance is typically compared through the comparison of sensitivity and specificity values or alternately by using another statistical comparison method such as receiver operator characteristic (ROC) curves or t-tests.

10.4.1 New Heuristic Development Without a Competing Heuristic

Development of a brand new heuristic decision model when another does not yet exist, may be seen as a special case of creating a new heuristic model that uses a significantly different set of variable elements for producing it's heuristic decisions. In this case, multiple neural network models are developed to determine the optimal variables for the heuristic and each model is compared against the others, with the currently best performing model serving as the existing heuristic model in Fig. 10.2.

Many of the earlier research projects listed in Table 10.1 fall into this class of a new heuristic model without any competing heuristic decision model (Baxt 1991, 1995; Baxt and Skora 1996; Frye et al. 1996; Izenberg et al. 1997). This is likely due to information processing for diagnostic purposes still being new to the medical field and heuristics to deal with the ever increasing quantity of patient information were only in early stages of development. These may also be viewed as proof of concept types of research that demonstrated a new tool, the neural network, for accurately evaluating large quantities of variables in a reliable manner, thus creating there own heuristic methodology.

Another motivation for development of brand new heuristic models is to satisfy the increasing pressure to reduce medical costs while maintaining or improving patient quality of care. The early prediction of trauma transfusion requirements NN utilizes information typically available at arrival in the emergency room (ER) without the need for invasive or costly laboratory tests (Walczak 2005). Mobley et al. (2005) demonstrate that a NN may accurately predict coronary disease using only variables available noninvasively. Neural networks have shown the ability to accurately diagnose the injury severity of pediatric trauma patients using only information available within the first 10 min of arrival in the ER (Walczak and Scorpio 2000). Finally, using a spectrometer and noninvasive data only, a neural network heuristic model has been able to accurately predict burn depth and severity (Yeong et al. 2005) This trend for developing diagnostic heuristics that utilize noninvasive test results or reduce the need for invasive procedure and thereby the risk to the patient is likely to continue as a driving goal in medical neural network research.

10.4.2 New Heuristic Development Competing Against an Existing Heuristic

The next classification of heuristic model development is for new heuristic models that compete against an existing heuristic model. Research for creating a new heuristic decision model that utilizes significantly different variables and also for modifying an existing decision heuristic to gain improved performance normally arises from dissatisfaction with the performance of the existing heuristic method or possibly simple intellectual curiosity to determine if any improvements are feasible.

Dissatisfaction with an existing heuristic, the MSBOS (Maximum Surgical Blood Order Supply) system, is the cause for the development of the transfusion prediction model that utilizes standard presurgery laboratory tests to significantly improve on blood unit ordering (Walczak and Scorpio 2000). As per Fig. 10.2, a neural network model utilizing physician determined variables was iteratively compared against the MSBOS across a large accumulation of historic transfusion data, with the neural network model consistently outperforming the old MSBOS heuristic when evaluated using a C/T ratio (units ordered/units transfused), by over 2 units average per patient. Refinements to the new heuristic were then accomplished using the leave-one-out technique described earlier to reduce the overall variable count for the neural network based heuristic transfusion resource model from 9 to 7 variables, which further decreased the C/T ratio by almost 0.7 units.

Another example, which actually disproves an existing heuristic method and recommends a new heuristic method came from a desire to see if neural networks were capable of improving predicting patient outcomes, in particular for acute pancreatitis (Pofahl et al. 1998; Walczak et al. 2003), over a more traditional method. In this neural network research to predict outcome severity with regard to hospital length of stay (LOS), numerous variables were identified and two sets of variables, with one being a subset of the first, were identified as predictive. Following the process in Fig. 10.2, each heuristic model was optimized and then compared against the other. The smaller variable set outperformed the larger set with evaluation being performed by comparing the mean absolute difference in LOS predicted by the neural network model vs. the actual LOS and percentage correct for 1–7, 8–14, and longer LOS. The larger variable set contained a commonly used heuristic variable, the abbreviated RANSON score, while the smaller set did not. Additional analysis was performed utilizing a regression model based on the abbreviated RANSON score as an alternate heuristic model. The abbreviated RANSON score was used, since a corollary focus of the research was on making the length of stay predictions utilizing information that was available within the first 12 h of presentation at an ER. The 22 variable LOS heuristic prediction model also outperformed the abbreviated RANSON-based regression model. This example raised questions about the current utilization of the RANSON scores as a valid heuristic for predicting acute pancreatitis patient outcomes.

The application of a new set of variables in medical decision heuristics may also result from advances in medical technology (Güler et al. 2005a), such as the inclusion of gene expression signatures for diagnosing breast cancer (Harbeck et al. 2007) and the combination of genetic variant variables to diagnose venous thrombosis (Penco et al. 2005). This is especially true with new imaging technologies, such as the utilization of Doppler velocity signals to accurately diagnose head trauma (Erol et al. 2005) and advances in blood imagining technology (Zini 2005). The research by Penco et al. (2005) also helps illustrate the new heuristic model refinement process of Fig. 10.2. After the new neural network heuristic model was convincingly shown to outperform existing heuristic models for diagnosing venous thrombosis, the new model was optimized by reducing the overall variable count from 62 values to 9 and finally down to just 3 variable values.

10.4.3 Improving an Existing Heuristic to Create a New Heuristic

The third heuristic development method seeks to improve the utilization and efficacy of existing heuristics instead of replacing them through the modification of the utilized variables in the heuristic model. This modification may be the addition of a new group of variables to a current set of variables or may involve the deletion of specific variables from an existing set of decision variables for a medical diagnostic or resource allocation heuristic. Deletion of highly correlated variables is necessary for neural networks models and will ultimately improve their classification or prediction performance.

A common theme in medicine, again, is reduction of costs and this may be achieved by reducing the unnecessary utilization of very expensive tests and treatments for patients that are falsely identified as belonging to a treatment group. Typically this is done by trying to improve the negative predictive capability of existing heuristics.

Several of the examples from Table 10.1 fall into this type of heuristic development, where improving the negative predictive capability of an existing heuristic is the goal of the research, with subsequent reduction in risk to patients and also a significant reduction in healthcare costs. Neural networks have been able to reduce false positives for: prostate cancer and subsequent reduction in unnecessary biopsies by being able to lower the PSA level at which a biopsy is ordered (Reckwitz et al. 1999), pulmonary embolism and reduction in chest radiography by adding in the reactive glucose variable with the existing d-dimer assay variable (Walczak et al. 2006), and earlier and more reliable detection of acute myocardial infarction by reducing the cutoff value for troponin l (Eggers et al. 2007). For each of these neural network heuristic diagnostic models incorporating the expanded variable set is compared, similar to Fig. 10.2, to either a separate neural network implementation or existing statistical model of the existing heuristic to demonstrate the efficacy of the new heuristic variable elements in improving diagnostic performance.

Another example of the heuristic refinement process from Fig. 10.2 for improving an existing heuristic is shown in the research by Güler and Übeyli (2005) who utilize a signal to noise ratio measurement to remove high noise variables from a neural network heuristic model to improve it's performance with a new smaller set of variable elements. As mentioned previously, the refinement of an existing heuristic through the removal of some variables further serves to reduce medical costs by reducing the data acquisition costs for the heuristic model (Bansal et al. 1993).

10.4.4 Methodological Heuristics

Each of the preceding medical domain examples has focused on the development of a better diagnostic or resource allocation heuristic method and involves modeling the new heuristic with a neural network implementation. A common claim made in the various research samples displayed in Table 10.1 is the efficacy and superiority of the neural network nonparametric modeling paradigm over more traditional statistical methods or domain expert performance. Much other neural network research is focused solely on demonstrating improvements to decision making through the utilization of neural network models. To distinguish these types of claims from the development of new heuristic models, this will be identified as improving the methodology or tool utilized to model the new problem solving heuristic, which may be seen as a fourth type of heuristic development where an existing heuristic model undergoes performance improvement without altering the variable elements of the heuristic model. These performance improvements are typically gained because the original model may have ignored some of the parametric requirements for the variables used in the heuristic method or a nonlinear interaction component between variable elements was present.

Academic research publications, especially in business, have failed to recognize the modeling power and statistical similarities of neural networks compared with other more commonly used parametric statistical models (e.g., regression and discriminant analysis). The equivalence of neural network modeling techniques compared to statistical modeling techniques has been shown in the literature for a wide variety of statistical techniques (Cheng and Titterington 1994; Raudys 1998; Zhang 2007), including: autoregression (Conner et al. 1994; Cottrell et al. 1995), canonical correlation analysis (Via et al. 2007), discriminant analysis (Gallinari et al. 1991), linear regression (Kumar 2005; Stern 1996; Warner and Misra 1996), logistic regression (Schumacher et al. 1996; Warner and Misra 1996), and maximum variance generalization (Via et al. 2007) among others.

As such, researchers frequently demonstrate the utility of their neural network models through comparison with traditional statistical models that utilize the same input variables. In Fig. 10.2 then, the existing heuristic would be a commonly used statistical model and the new heuristic would be the same model implemented as a neural network. It is important though to when comparing a neural network diagnostic or classification heuristic model against existing statistical models that an appropriate model is selected. Commonly used statistical models in medical domains include stepwise linear regression, logistic regression, and discriminant analysis (Walczak and Scorpio 2000; Zhang and Berardi 1998).

Neural networks have been shown to outperform both parametric and nonparametric statistical models across a wide variety of domains. Examples of superior neural network performance over traditional statistical methods are: in medical domains (Baxt and Skora 1996; Dybowski and Gant 1995; Lapuerta et al. 1995, León 1994; Razi and Athappilly 2005; Rodvold et al. 2001; Zhang and Berardi 1998) and business domains (Bansal et al. 1993; Devika and Achenie 1994; Falas et al. 1994; Lee et al. 1993; Piramuthu et al. 1994; Refenes 1993; Steurer 1993). These results were foretold by Patuwo et al. (1993) who demonstrate that using more sophisticated modeling techniques like neural networks can improve heuristic classification model performance by 15–22%.

The resulting heuristic-oriented research simply tries to improve upon the speed of the availability of diagnostic or resource information or improve performance without modifying the set of variables utilized by the current heuristic methods. Frequently this type of research may also argue for the utilization of neural network training methods other than the traditional multilayer perceptron backpropagation method. Examples of methodological-oriented heuristic improvement through translating an existing decision model into a neural network representation are: The use of cellular neural networks are proposed tom rapidly improve the speed of image processing for diagnosis of cancer (Arena et al. 2003), the use of probabilistic neural network models to improve performance in two separate medical diagnostic problems and also two protein localization problems (Georgiou et al. 2006), neural networks that automate the detection of metastases in bone scans with very high sensitivity and specificity (Sadik et al. 2006), a neural network model produced an 8% improvement in sensitivity for predicting ischemic heart disease (Scott et al. 2004), accuracy for predicting carotid artery stenosis was improved through use of a neuron-fuzzy system (Übeyli and Güler 2005a), and a mixture of experts (multiple) neural network system was able to obtain an almost 99% accuracy performance for diagnosing breast cancer (Übeyli and Güler 2005b). In addition to a neural network application providing the desired improvements, this type of research may also produce recommendations for combining neural network heuristic models with other applications, such as the combination of a neural network with a decision tree model to more accurately predict cancer relapse (Jerez-Aragones et al. 2003) and a combination of a neural network to refine EEG tracings with an expert system (Castellaro et al. 2002).

10.5 Neural Network Heuristic Evaluation in Business Domains

Businesses outside of the medical domain also face complex decision making problems where heuristic techniques are applied to simplify or speed up the decision making process. The process for evaluating heuristics in more general business domains is identical to the heuristic evaluation method previously described and shown in medical domains.

This section will start by listing business-oriented neural network research that specifically claims to create new heuristics or improve upon existing heuristics, similar to how neural network heuristic models were shown in the medical heuristics section. Table 10.2 lists neural network heuristic research in business domains.

Because of the frequent application of neural networks in solving business and engineering decision problems, some interesting aspects of utilizing neural networks for heuristic development emerge in the business domains. Both statistics and neural networks have been the primary methods for evaluating credit card applications and risks (He et al. 2004) and as such the neural network models may serve as the existing model in Fig. 10.2 for comparison to new methods including new neural network techniques. Besides serving as the existing heuristic model for new research, neural networks have also been proposed as an efficient methodology for selecting between existing and competing heuristics (Gupta et al. 2000), which would mean that the neural network is serving as a heuristic method to select the most appropriate decision heuristic, which could potentially include other neural network heuristic models.

A few select neural network heuristic development cases from Table 10.2 are now analyzed in greater detail to demonstrate the neural network heuristic comparison methodology shown in Fig. 10.2. The first two examples represents the evaluation using neural networks of multiple existing heuristics that compete within a domain. The competing heuristics are: utilization of technical vs. fundamental analysis models, the financial heuristic that "more

Table 10.2. Samples of neural network (NN) heuristic research in business

Business domain	Heuristic development type[a]	Heuristic result	Citation
Accounting	3 (or 4)	Establishes bridge to economic accelerator models	(Fioretti 2004)
	2	NN used to develop new methodology for creating bankruptcy prediction models recommending representative samples vs. stratified samples for model training	(Wilson and Sharda 1994)
Finance	4	Uses PCA component analysis to improve NN predictions	(Ince and Trafalis 2007)
	4	NN demonstrates futility of day trading stable stocks	(Tsang et al. 2007)
	3	NN validates global knowledge needed for predicting emerging market financial time series	(Walczak 1999, 2004)
	2	NN shows minimal knowledge needed to predict financial time series	(Walczak 2001)
	3	NN reduces loan default for lender by 10%	(West 2000)
Human Resources	2	NN improves scheduling of airline ground staff	(Hao et al. 2004)
	1	NN utilized to predict enrollment patterns to schedule admissions personnel	(Walczak 1998; Walczak and Sincich 1999)
Management	2	NN used to optimize load allocations for cogeneration power plants	(Cerri et al. 2006)
	3	NN optimizes nuclear fuel management	(Jiang et al. 2006)
Manufacturing	4 (or 3)	NN combines greedy and nongreedy heuristics to improve job scheduling	(Agarwal et al. 2006)
	4	NN improves flow shop scheduling using an existing heuristic	(Haq and Ramanan 2006)
	4	Combination of NN and decision trees reduce false classification of control chart patterns	(Ruey-Shiang 2005)

Table 10.2. (continued)

Business domain	Heuristic development type[a]	Heuristic result	Citation
	2	NN improves on real-time job shop scheduling	(Shugang et al. 2005)
	2	NN minimizes exceptional elements through fractional cell formation in group technology	(Venkumar and Haq 2006)
	2	NN combines flow shop and parallel machine algorithms to produce new heuristic for flexible flow shop scheduling	(Wang et al. 2003)
Marketing	3	Determining heuristic variables with independent component analysis produces NN that outperforms all other statistical methods	(Ahn et al. 2007)
	1	NN uses navigation patterns to accurately predict customer product knowledge, if combined with more traditional surveys, even greater improvement is possible	(Chang et al. 2006)
	4	NN significantly improves service discontinuation predictions over multiple discriminant analysis models	(Walczak and Parthasarathy 2006)
Transportation	2	NN improves the responsiveness and reliability of paratransit systems	(Fu and Teply 1999)
	4	NN improves traffic predictions over other heuristic methods	(Ishak and Alecsandru 2004)
	4	NN as a part of an Expert System maximizes air freight shipping	(Lau et al. 2004)

[a] 1 = New heuristic w/o a competing heuristic, 2 = new heuristic w/competing heuristic, 3 = modification of existing heuristic, 4 = methodological improvement of existing heuristic

information produces better models", and the development of bankruptcy prediction models using representative or stratified data sets.

Neural network models, with the addition of some variables, that represent both technical and fundamental models are implemented and evaluated against each other. The various fundamental and technical models implemented in neural networks serve as the two heuristic models for comparison in Fig. 10.2. Since the two models are competing on an equal basis, with neither one considered.

The existing model, refinement of the independent variables for each model is performed simultaneously and reevaluation of all new models is performed. The model selection research methodology (Swanson and White 1995, 1997) is used to select the best performing neural network heuristic model, with the evaluation for both competing models being performed across the same data set for the dependent variables.

Neural network financial time series, especially in foreign exchange rate predictions, typically utilize a technical and homogeneous model, utilizing a single lag[1] equal to the time period being forecast (e.g., one day lags for a single day forecast or a five-day lag for a one week forecast). Neural network research has shown that using a more fundamental analysis type of model, which includes values for macroeconomic variables outside of the exchange rate itself would improve forecasting capabilities of neural network heuristic exchange rate prediction models over a more technical model (Walczak 2001). A corollary outcome of this research showed that financial time series are cyclical and that training forecasting models with data beyond the first full cycle was unnecessary, which contradicted the existing financial ideal of "the more data the better".

This heuristic evaluation result of fundamental or heterogeneous models outperforming technical or homogeneous models was further exploited in research that demonstrated how global financial indicators are required of heuristic models attempting to forecast stock market index futures in emerging markets in the Pacific Rim (Walczak 1999) and South America (Walczak 2004). In each case, both technical and fundamental analysis heuristic models were developed and evaluated against each other, with the fundamental heuristic models consistently demonstrating superior forecasting performance and the forecasting performance exceeding traditional random walk models of financial forecasting. The fund

However, technical models still tend to dominate neural network and information systems models of financial markets in general. This is because the models themselves are simpler due to the reliance on only a single variable or homogenous set of variables derived from the original forecasting value, which reduces variable costs (Bansal et al. 1993) and ultimately makes the

[1] A lag is the difference between a financial value at one time, typically the day prior to the forecast, and another past day. A one day lag for a financial value v at time t is $v_t - v_{t-1}$ and a five day lag for the same value at time t is $v_t - v_{t-5}$.

heuristic model's results easier to interpret. Other research has focused on the second heuristic development method described in Sect. 4.4, the improvement of an existing heuristic through the addition or deletion of variables, to maximize the performance of these technical models. The reported research (Walczak 2001; Walczak et al. 1998) demonstrates that utilizing multiple lag values significantly improves the forecasting ability of these neural network heuristic models. Again, it should be emphasized that the evaluation of these heuristics is done by statistically analyzing the forecasting performance of the single day lag neural network model against other neural network models utilizing multiple lag values in the input vector set and also statistical models (e.g., ARIMA).

The final example concerns the propensity for bankruptcy models to be constructed utilizing stratified data sets (typically a 50–50 distribution). Early research indicated that heuristic bankruptcy prediction neural network models could be optimized using representative training data, meaning that the training data reflected the group proportionality of the real world (Hu et al. 1996; Wilson and Sharda 1994). These neural network models evaluating the possibility of representative training samples were evaluated against similar neural network models using various stratified training samples and also against statistical models (e.g., logistic regression) also with representative and stratified samples. All results for evaluation were performed against a representative data sample to reflect utilization of these heuristic models This type of finding would make data collection for model building easier since it could be done automatically without performing any matching to stratify the training (model development) data sets. However, other research has more recently contradicted these findings, indicating that neural networks and statistical models need to be trained on a 50/50 stratified sample to optimize performance by keeping the neural network from becoming trapped in a local minimum (Sharda and Wilson 1996). Again, the more recent research duplicated the methodology of simultaneously comparing multiple neural network and statistical models developed using both representative and stratified training samples, only with a different data set.

Like the medical heuristic development through neural networks, the examples described in this section and given in Table 10.2 demonstrate that neural networks are a reliable and effective mechanism for evaluating and developing business related heuristics. However, as the last case has pointed out, car must be taken when generalizing results and new heuristics should be validated across multiple data sets before ultimately replacing an existing heuristic.

10.6 Conclusions

Neural networks are commonly used in business (Smith and Gupta 2000; Wong et al. 2000) and medicine (Baxt 1995; Dybowski and Gant 1995; Zini 2005). Typically, new neural network applications implement heuristic methods for new problems or to improve performance over existing decision heuristics. Neural networks, because of their nonparametric nature serve as an outstanding evaluation tool for comparison of heuristic models and evaluation of new heuristic models. Lewin (2000) goes so far as to state that neural networks are a cross-over methodology between artificial intelligence and regression.

This chapter has defined a more formal process for utilizing neural networks in heuristic decision model evaluation. Several examples of how the process has been applied in previous research in both medical domains and business domains have been presented to demonstrate the application of the described process. The process is meant to serve as a guide to researchers and professional developers for evaluating competing heuristic decision models utilizing neural networks. One or both of the competing models may be implemented as a neural network, but comparison against traditional nonneural network heuristic and statistical models is applicable.

Feng et al. (2006) examine the development and evaluation of a regression and neural network model in the manufacturing business domain. There process of developing two competing models, based on existing heuristic methods, and then evaluating them statistically to compare the results and select the optimal model is similar to the proposed methodology formalism in this chapter. Ultimately in their research, the neural network and regression models were shown to be equivalent and the new heuristic model (which is actually what this chapter calls a methodological heuristic advancement) implementation selection is based on other factors, which in this case is the belief that neural networks are a black-box methodology.

Researchers must realize that neural network models are as rigorous as more traditional statistical models and have distinct advantages for modeling complex and potentially nonlinear business and medical decision making heuristics, such as tolerance of noise in the data, fast training and real-time results, and nonparametric capability to model numerous population and error distributions.

References

Abdel-Aal, R.E. (2004) Experimental evaluation of performance improvements in abductive network classifiers with problem decomposition. *Neurocomputing* 61(1–4), 193–215

Agarwal, A., Jacob, V.S., and Pirkul, H. (2006) An improved augmented neural-network approach for scheduling problems. *INFORMS Journal on Computing* 18(1), 119–128

Ahn, H., Choi, E., and Han, I. (2007) Extracting underlying meaningful features and canceling noise using independent component analysis for direct marketing. *Expert Systems with Applications* 33(1), 181–191

Arena, P., Basile, A., Bucolo, M., and Fortuna, L. (2003) Image processing for medical diagnosis using CNN. *Nuclear Instruments and Methods in Physics Research Section A* 497(1), 174–178

Ballini, R. and Gomide, F. (2002/2003) Heuristic learning in recurrent neural fuzzy networks. *Journal of Intelligent and Fuzzy Systems* 13(2–4), 63–74

Bansal, A., Kauffman, R.J., and Weitz, R.R. (1993) Comparing the modeling performance of regression and neural networks as data quality varies: A business value approach. *Journal of Management Information Systems* 10(1), 11–32

Baxt, W.G. (1991) Use of an artificial neural network for the diagnosis of myocardial infarction. *Annals of Internal Medicine* 115, 843–848

Baxt, W.G. (1995) Application of artificial neural networks to clinical medicine. *The Lancet* 346(8983), 1135–1138

Baxt, W.G. and Skora, J. (1996) Prospective validation of artificial neural network trained to identify acute myocardial infarction. *The Lancet* 347, 12–15

Buchman, T.G., Kubos, K.L., Seidler, A.J., and Siegforth, M.J. (1994) A comparison of statistical and connectionist models for the prediction of chronicity in a surgical intensive care unit. *Critical Care Medicine* 22(5), 750–762

Castellaro, C., Favaro, G., Castellaro, A., Casagrande, A., Castellaro, S., Puthenparampil, D.V., and Fattorello Salimbeni, C. (2002) An artificial intelligence approach to classify and analyse EEG traces. *Clinical Neurophysiology* 32(3), 193–212

Cerri, G., Borghetti, S., and Salvini, C. (2006) New management for heat and power cogeneration plants. *Engineering Applications of Artificial Intelligence* 19(7), 721–730

Chang, S.E., Changchien, S.W., and Huang, R.-H. (2006) Assessing users' product-specific knowledge for personalization in electronic commerce. *Expert Systems with Applications* 30(4), 682–693

Cheng, B. and Titterington, D.M. (1994) Neural networks: A review from a statistical perspective. *Statistical Science* 9(1), 2–30

Connor, J.T., Martin, R.D., and Atlas, L.E. (1994) Recurrent neural networks and robust time series prediction. *IEEE Transactions on Neural Networks* 51(2), 240–254

Coppini, G., Miniati, M., Paterni, M., Monti, S., and Ferdeghini, E.M. (2007) Computer-aided diagnosis of emphysema in COPD patients: Neural-network-based analysis of lung shape in digital chest radiographs. *Medical Engineering and Physics* 29(1), 76–86

Cottrell, M., Girard, B., Girard, Y., Mangeas, M., and Muller, C. (1995) Neural modeling for time series: A statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks* 6(6), 1355–1364

Daniel, W.W. (1999) *Biostatistics: A Foundation for Analysis in the Health Sciences*. New York: Wiley

Devika, P.D. and Achenie, L.E.K. (1994) Experience with Using a Neural Network for Forecasting Time Series. In: *World Congress on Neural Networks – San Diego Volume 2*, San Diego, CA, pp. 315–320

Dybowski, R. and Gant, V. (1995) Artificial neural networks in pathology and medical laboratories. *The Lancet* 346, 1203–1207

Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Siam

Eggers, K.M., Ellenius, J., Dellborg, M., Groth, T., Oldgren, J., Swahn, E., and Lindahl, B. (2007) Artificial neural network algorithms for early diagnosis of acute myocardial infarction and prediction of infarct size in chest pain patients. *International Journal of Cardiology* 114(3), 366–374

Erol, F.S., Uysal, H., Ergün, U., Barişçi, N., Serhathoğlu, S., and Hardalaç, F. (2005) Prediction of minor head injured patients using logistic regression and MLP neural network. *Journal of Medical Systems* 29(3), 205–215

Falas, T., Charitou, A., and Charalambous, C. (1994) The Application of Artificial Neural Networks in the Prediction of Earnings. In: *Proceedings of IEEE International Conference on Neural Networks Volume VI*, Orlando, FL, pp. 3629–3633

Feng, C.-X.J., Yu, Z.-G., Kusiak, A. (2006) Selection and validation of predictive regression and neural network models based on designed experiments. *IIE Transactions* 38(1), 13–23

Fioretti, G. (2004) The investment acceleration principle revisited by means of a neural network. *Neural Computing and Applications* 13(1), 16–23

Frye, K.E., Izenberg, S.D., Williams, M.D., and Luterman, A. (1996) Simulated biologic intelligence used to predict length of stay and survival of burns. *Jounral of Burn Care and Rehabilitation* 17(6), 540–546

Fu, L. and Teply, S. (1999) On-line and off-line routing and scheduling of dial-a-ride paratransit vehicles. *Computer-Aided Civil and Infrastructure Engineering* 14(5), 309–319

Gallinari, P., Thiria, S., Badran, R., and Fogelman-Soulie, F. (1991) On the relationships between discriminant analysis and multilayer perceptrons. *Neural Networks* 4(3), 349–360

Georgiou, V.L., Pavlidis, N.G., Parsopoulos, K.E., Alevizos, P.D., and Vrahatis, M.N. (2006) New self-adaptive probabilistic neural networks in bioinformatic and medical tasks. *International Journal of Artificial Intelligence Tools* 15(3), 371–396

Gjertsson, P., Lomsky, M., Richter, J., Ohlsson, M., Tout, D., van Aswegen, A., Underwood, R., and Edenbrandt, L. (2006) The added value of ECG-gating for the diagnosis of myocardial infarction using myocardial perfusion scintigraphy and artificial neural networks. *Clinical Physiology and Functional Imaging* 26(5), 301–304

Güler, İ., Polat, H., and Ergün, U. (2005a) Combining neural network and genetic algorithm for prediction of lung sounds. *Journal of Medical Systems* 29(3), 217–231

Güler, İ. and Übeyli, E.D. (2005) Feature saliency using signal-to-noise ratios in automated diagnostic systems developed for ECG beats. *Expert Systems with Applications* 28(2), 295–304

Güler, N.F., Übeyli, E.D., and Güler, İ. (2005b) Recurrent neural networks employing Lyapunov exponents for EEG signals classification. *Expert Systems with Applications* 29(3), 506–514

Gupta, J.N.D., Sexton, R.S., and Tunc, E.A. (2000) Selecting scheduling heuristics using neural networks. *INFORMS Journal on Computing* 12(2), 150–162

Hao, G., Lai, K.K., and Tan, M. (2004) A neural network application in personnel scheduling. *Annals of Operations Research* 128(1–4), 65–90

Haq, A.N. and Ramanan, T.R. (2006) A bicriterian flow shop scheduling using artificial neural network. *International Journal of Advanced Manufacturing Technology* 30(11/12), 1132–1138

Harbeck, N., Schmitt, M., Paepke, S., Allgayer, H., and Kates, R.E. (2007) Tumor-associated proteolytic factors uPA and PAI-1: Critical appraisal of their clinical relevance in breast cancer and their integration into decision-support algorithms. *Critical Reviews in Clinical Laboratory Sciences* 44(2), 179–201

He, J., Liu, X., Shi, Y., Xu, W., and Yan, N. (2004) Classification of credit cardholder behavior by using Fuzzy linear programming. *International Journal of Information Technology and Decision Making* 3(4), 633–650

Hertz, J., Krogh, A., and Palmer, R.G. (1991) *Introduction to the Theory of Neural Computation.* Reading, MA: Addison-Wesley

Hu, M.Y., Hung, M.S., Shanker, M.S., and Chen, H. (1996) Using neural networks to predict performance of Sino-foreign joint ventures. *International Journal of Computational Intelligence and Organizations* 1(3), 134–143

Ince, H. and Trafalis, T.B. (2007) Kernel principal component analysis and support vector machines fro stock price prediction. *IIE Transactions* 39(6), 629–637

Ishak, S. and Alecsandru, C. (2004) Optimizing traffic prediction performance of neural networks under various topological, input, and traffic condition settings. *Journal of Transportation Engineering* 130(4), 452–465

Izenberg, S.D., Williams, M.D., and Luterman, A. (1997) Prediction of trama mortality using a neural network. *American Surgeon* 63(3), 275–281

Jain, A.K., Mao, J., and Mohiuddin, K.M. (1996) Artificial neural networks: A tutorial. *IEEE Computer* 29(3), 31–44

Jerez-Aragones, J.M., Gomez-Ruiz, J.A., Ramos-Jimenez, G., Munoz-Perez, J., and Albaconejo, E. (2003) A combined neural network and decision trees model for prognosis of breast cancer relapse. *Artificial Intelligence in Medicine* 27(1), 45–63

Jiang, S., Ziver, A.K., Carter, J.N., Pain, C.C., Goddard, A.J.H., Franklin, S., and Phillips, H.J. (2006) Estimation of distribution algorithms for nuclear reactor fuel management optimization. *Annals of Nuclear Energy* 33(11/12), 1039–1057

Kumar, U.A. (2005) Comparison of neural networks and regression analysis: A new insight. *Expert Systems with Applications* 29(2), 424–430

Lapuerta, P., Azen, S.P., and LaBree, L. (1995) Use of neural networks in predicting the risk of coronary artery disease. *Computers in Biomedical Research* 28(1), 38–52

Lau, H.C.W., Choy, K.L., Lau, P.K.H., Tsui, W.T., and Choy, L.C. (2004) An intelligent logistics support system for enhancing the airfreight forwarding business. *Expert Systems* 21(5), 253–268

Lee, T., White, H., and Granger, C. (1993) Testing for neglected nonlinearity in time series models. *Journal of Econometrics* 56(3), 269–290

León, M.A. (1994) Binary response forecasting: comparison between neural networks and logistic regression analysis, In: *World Congress on Neural Networks*, San Diego, CA, pp. 244–247

Lewin, D.I. (2000) Getting clinical about neural networks. *IEEE Intelligent Systems and Their Applications* 15(1), 2–5

Lim, T., Loh, W., and Yu-Shan, S. (2000) A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning* 40(3), 203–228

Lisboa, P.J.G. (2002) A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Networks* 15(1), 11–39

Lisboa, P.J. and Taktak, A.F.G. (2006) The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* 19(4), 408–415

Markey, M.K., Tourassi, G.D., Margolis, M., and DeLong, D.M. (2006) Impact of missing data in evaluating artificial neural networks trained on complete data. *Computers in Biology and Medicine* 36(5), 516–525

McLachlan, G.J. (1992) *Discriminant Analysis and Statistical Pattern Recognition*. New York: Wiley

Mobley, B.A., Schechter, E., Moore, W.E., McKee, P.A., and Eichner, J.E. (2005) Neural network predictions of significant coronary artery stenosis in men. *Artificial Intelligence in Medicine* 34(2), 151–161

Nath, R., Rajagopalan, B., and Ryker, R. (1997) Determining the saliency of input variables in neural network classifiers. *Journal of Computers and Operations Research* 24(8), 767–773

Nowack, W.J., Walczak, S., and Janati, A. (2002) Clinical correlate of EEG rhythmicity. *Journal of Clinical Neurophysiology* 19(1), 32–36

Pakath, R. and Zaveri, J. S. (1995) Specifying Critical Inputs in a Genetic Algorithm-driven Decision Support System: An Automated Facility. *Decision Sciences* 26(6), 749–779

Patuwo, E., Hu, M.Y., and Hung, M.S. (1993) Two-group classification using neural networks. *Decision Sciences* 24(4), 825–845

Pearl, J. (1984) *Heuristics – Intelligent Search Strategies for Computer Problem Solving*. Reading, MA: Addison-Wesley

Penco, S., Grossi, E., Cheng, S., Intraligi, M., Maurelli, G., Patrosso, M.C., Marocchi, A., and Buscema, M. (2005) Assessment of the Role of Genetic Polymorphism in Venous Thrombosis Through Artificial Neural Networks. *Annals of Human Genetics* 69(6), 693–706

Piramuthu, S., Shaw, M., and Gentry, J. (1994) A classification approach using multi-layered neural networks. *Decision Support Systems* 11(5), 509–525

Pofahl, W.P., Walczak, S.M., Rhone, E., and Izenberg, S.D. (1998) Use of an artificial neural network to predict length of stay in acute pancreatitis. *American Surgeon* 64(9), 868–872

Raudys, S. (1998) Evolution and generalization of a single neuron – I: Singlelayer perceptron as seven statistical classifiers. *Neural Networks* 11(2), 283–296

Razi, M.A. and Athappilly, K. (2005) A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Systems with Applications* 29(1), 65–74

Reckwitz, T., Potter, S.R., Snow, P.B., Zhang, Z., Veltri, R.W., and Partin, A.W. (1999) Artificial neural networks in urology: Update 2000. *Prostate Cancer and Prostatic Diseases* 2(5/6), 222–226

Refenes, A.N. (1993) Constructive Learning and Its Application to Currency Exchange Rate Forecasting. In: R. Trippi and E. Turban (Eds.), *Neural Networks in Finance and Investing*, New York: Irwin, pp. 465–493

Rodvold, D.M., McLeod, D.G., Brandt, J.M., Snow, P.B., and Murphy, G.P. (2001) Introduction to artificial neural networks for physicians: Taking the lid off the black box. *The Prostate* 46(1), 39–44

Ruey-Shiang, G. (2005) Real-time pattern recognition in statistical process control: a hybrid neural network/decision tree-based approach. *Proceedings of the Institution of Mechanical Engineers – Part B – Engineering Manufacture* 219(3), 283–298

Sadik, M., Jakobsson, D., Olofsson, F., Ohlsson, M., Suurkula, M., and Edenbrandt, L. (2006) A new computer based decision support system for the interpretation of bone scans. *Nuclear Medicine Communications* 27(5), 417–423

Salamalekis, E., Siristatidis, C., Vasios, G., Saloum, J., Giannaris, D., Chrelias, C., Prentza, A., and Koutsouris, D. (2006) Fetal pulse oximetry and wavelet analysis of the fetal heart rate in the evaluation of abnormal cardiotocography tracings. *Journal of Obstetrics and Gynaecology Research* 32(2), 135–139

Schumacher, M., Robner, R., and Vach, W. (1996) Neural networks and logistic regression – Part I. *Computational Statistics and Data Analysis* 21(6), 661–682

Schürmann, J. (1996) *Pattern Classification a Unified View of Statistical and Neural Approaches*. New York: Wiley

Scott, J.A., Aziz, K., Yasuda, T., and Gewirtz, H. (2004) Integration of clinical and imaging data to predict the presence of coronary artery disease with the use of neural networks. *Coronary Artery Disease* 15(7), 427–434

Sharda, R. and Wilson, R.L. (1996) Neural network experiments in business-failure forecasting: Predictive performance measurement issues. *International Journal of Computational Intelligence and Organizations* 1(2), 107–117

Shin, K.S., Lee, T.S., and Kim, H. (2005) An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications* 28(1), 127–135

Shugang, L., Zhiming, W., and Xiaohong, P. (2005) Job shop scheduling in real-time cases. *International Journal of Advanced Manufacturing Technology* 26(7/8), 870–875

Smith, M. (1993) *Neural Networks for Statistical Modeling*, New York: Van Nostrand Reinhold

Smith, K.A. and Gupta, J.N.D. (2000) Neural networks in business: techniques and applications for the operations researcher. *Computers and Operations Research* 27(11–12), 1023–1044

Srinivasan, V., Eswaran, C., and Sriraam, N. (2005) Artificial Neural Network Based Epileptic Detection Using Time-Domain and Frequency-Domain Features. *Journal of Medical Systems* 29(6), 647–660

Soulié, F. (1994) Integrating Neural Networks for Real World Applications. In: Zurada, Marks, and Robinson (Eds.), *Computational Intelligence: Imitating Life*, New York: IEEE Press, pp. 396–405

Stephan, C., Jung, K., Cammann, H., Vogel, B., Brux, B., Kristiansen, G., Rudolph, B., Hauptmann, S., Lein, M., Schnorr, D., Sinha, P., and Loening, S.A. (2002) An artificial neural network considerably improves the diagnostic power of percent free prostate-specific antigen in prostate cancer diagnosis: results of a 5-year investigation. *International Journal of Cancer* 99(3), 466–473

Stern, H.L. (1996) Neural networks in applied statistics. *Technometrics* 38(3), 205–214

Steurer, E. (1993) Nonlinear Modelling of the DEM/USD Exchange Rate. In: A. Refenes (Ed.), *Neural Networks in the Capital Markets*, New York: Wiley, pp. 199–211

Swanson, N.R. and White, H. (1995) A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks. *Journal of Business and Economic Statistics* 13(3), 265–275

Swanson N.R. and White, H. (1997) A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics* 79(4), 540–550

Tahai, A., Walczak, S., and Rigsby, J.T. (1998) Improving Artificial Neural Network Performance Through Input Variable Selection. In: P. Siegel, K. Omer, A. deKorvin, and A. Zebda (Eds.) *Applications of Fuzzy Sets and the Theory of Evidence to Accounting II*, Stamford, Connecticut: JAI Press, pp. 277–292

Tourassi, G.D. and Floyd, C.E. (1997) The effect of data sampling on the performance evaluation of artificial neural networks in medical diagnosis. *Medical Decision Making* 17(2), 186–192

Tsang, P.M., Kwok, P., Choy, S.O., Kwan, R., Ng, S.C., Mak, J., Tsang, J., and Koong, K. (2007) Design and implementation of NN5 for Hong Kong stock price forecasting. *Engineering Applications of Artificial Intelligence* 20(4), 453–461

Übeyli, E.D. and Güler, I. (2005a) Adaptive neuro-fuzzy inference systems for analysis of internal carotid arterial Doppler signals. *Computers in Biology and Medicine* 35(8), 687–702

Übeyli, E.D. and Güler, I. (2005b) A mixture of experts network structure for breast cancer diagnosis. *Journal of Medical Systems* 29(5), 569–579

Venkumar, P. and Haq, A.N. (2006) Fractional cell formation in group technology using modified ART1 neural networks. *International Journal of Advanced Manufacturing Technology* 30(11/12), 1132–1138

Via, J., Santamaría, I., and Pérez, J. (2007) A learning algorithm for adaptive canonical correlation analysis of several data sets. *Neural Networks* 20(1), 139–152

Walczak, S. (1998) Neural network models for a resource allocation problem. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics* 28(2), 276–284

Walczak (1999) Gaining competitive advantage for trading in emerging capital markets with neural networks. *Journal of Management Information Systems* 16(2), 177–192

Walczak, S. (2001) An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of Management Information Systems* 17(4), 203–222

Walczak, S. (2004) Forecasting Emerging Market Indexes with Neural Networks. In: G.P. Zhang (Ed.), *Neural Networks in Business Forecasting*, Hershey, PA: Idea Group Publishing, pp. 80–101

Walczak, S. (2005) Artificial neural network medical decision support tool: Predicting Transfusion Requirements of ER Patients. *IEEE Transactions on Information Technology in Biomedicine* 9(3), 468–474

Walczak, S., Brimhall, B.B., and Lefkowitz, J.B. (2006) Nonparametric decision support systems in medical diagnosis: Modeling pulmonary embolism. *International Journal of Healthcare Information Systems and Informatics* 1(2), 65–82

Walczak, S. and Nowack, W.J. (2001) An artificial neural network approach to diagnosing epilepsy using lateralized bursts of theta EEGs. *Journal of Medical Systems* 25(1), 9–20

Walczak, S. and Parthasarathy, M. (2006) Modeling online service discontinuation with nonparametric agents. *Information Systems and e-Business Management* 4(1), 49–70

Walczak, S., Pofahl, W.E., and Scorpio, R.J. (2003) A decision support tool for allocating hospital bed resources and determining required acuity of care. *Decision Support Systems* 34(4), 445–456

Walczak, S. and Scharf, J.E. (2000) Reducing surgical patient costs through use of an artificial neural network to predict transfusion requirements. *Decision Support Systems* 30(2), 125–138

Walczak, S. and Scorpio, R.J. (2000) Predicting pediatric length of stay and acuity of care in the first ten minutes with artificial neural networks. *Pediatric Critical Care Medicine* 1(1), 42–47

Walczak, S. and Sincich, T. (1999) A comparative analysis of regression and neural networks for university admissions. *Information Sciences* 119(1–2), 1–20

Walczak, S., Tahai, A., and Khondar, K. (1998) Improved Cash Flows Using Neural Network Models for Forecasting Foreign Exchange Rates. In: P. Siegel, K. Omer, A. deKorvin, and A. Zebda (Eds.) *Applications of Fuzzy Sets and The Theory of Evidence to Accounting II*, Stamford, Connecticut: JAI Press, pp. 293–310

Wang, H., Jacob, V., and Rolland, E. (2003) Design of efficient hybrid neural networks fro flexible flow shop scheduling. *Expert Systems* 20(4), 208–231

Warner, B. and Misra, M. (1996) Understanding neural networks as statistical tools. *The American Statistician* 50(4), 284–293

Weigand, A.S. and Zimmermann, H.G. (1995) The Observer-Observation Dilemma in Neuro-Forecasting: Reliable Models From Unreliable Data Through CLEARNING. *Proceedings of Artificial Intelligence Applications on Wall Street*, New York, pp. 308–317

West, D. (2000) Neural Network Credit Scoring Models. *Computers and Operations Research* 27(11/12), 1131–1152

Wilson, R.L. and Sharda, R. (1994) Bankruptcy prediction using neural networks. *Decision Support Systems* 11(5), 545–557

Wong, B.K., Lai, V.S., and Lam, J. (2000) A bibliography of neural network business applications research: 1994–1998. *Computers and Operations Research* 27(11–12), 1045–1076

Yeong, E.-K., Hsiao, T.-C., Chiang, H.K., and Lin, C.-W. (2005) Prediction of burn healing time using artificial neural networks and reflectance spectrometer. *Burns* 31(4), 415–420

Zahedi, F. (1996) A meta-analysis of financial applications of neural networks. *International Journal of Computational Intelligence and Organizations* 1(3), 164–178

Zhang, G.P. (2007) Avoiding pitfalls in neural network research. *IEEE Transactions on Systems, Man, and Cybernetics – Part C* 37(1), 3–16

Zhang, G.P. and Berardi, V.L. (1998) An investigation of neural networks in thyroid function diagnosis. *Health Care Management Science* 1(1), 29–37

Zhang, G.P. and Hu, M.Y. (1998) Neural network forecasting of the British pound/US dollar exchange rate. *Omega, International Journal of Management Science* 26(4), 495–506

Zini, G. (2005) Artificial intelligence in Hematology. *Hematology* 10(5), 393–400