INTERNET NEGOTIATION PATTERNS AND AN APPRAISAL OF DATA MINING FROM A MANAGERIAL PERSPECTIVE: A CASE STUDY APPROACH

Yiwei Zhang

A Thesis
In
The John Molson School of Business

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Administration at
Concordia University
Montreal, Quebec, Canada

September 2001

0-612-64054-X

ABSTRACT

Internet Negotiation Patterns and an Appraisal of Data Mining from a Managerial Perspective: A Case Study Approach

Yiwei Zhang

Negotiation on the Internet is a new business activity that emerged with the development of the Internet and the World Wide Web. In order to study the use of software tools in cross-cultural Internet negotiations, a project named InterNeg was initiated. The INSPIRE negotiation support system of this project collected data of Internet negotiations by allowing participants to negotiate a mock case.

Empirical research was conducted on these data by applying different data mining methods. This is because there were few former studies and hypotheses, and the variable number is large and they are not obviously correlated. Three data mining methods were applied to find hidden behavior patterns of Internet negotiations: Tree Rule Induction (TRI), Artificial Neural Networks (ANN) and Logistic Regression Analysis (LRA). The results showed that the numbers of offers sent, especially during the early and middle negotiation stages, are positively related to reaching agreements, while sending offers at last minute has low chance to get the compromise. Other factors, such as gender and interval time between offers, can also affect Internet negotiation results.

Comparisons of results from different data mining, especially on their prediction accuracies, were also conducted. The results revealed that the TRI method enjoys the highest prediction accuracy while consuming least processing time. The ANN method has the lowest prediction accuracy. Our research results also indicated that the layer number and hidden unit number in the layers could not affect the ANN method's prediction accuracy.

Table of Contents

1 Introduction

Negotiation, as a social behavior, has been widely studied in recent decades from many perspectives. With the widespread use of networked computers, studies were also conducted on the use of computing tools in conducting and supporting negotiations (Chatterjee and Lilien 1984; Bui 1994). In such negotiations, the parties are aware of the identities of each other, and they usually use tools to support their face-to-face bargaining (Mahadevan 1999).

Recent advancements in network technology - especially the use of Email, Internet and the World Wide Web - have accelerated the development of net-centric tools (Carroll 1996). These tools allow their users to communicate with one another, even anonymously. In mid-nineties, Internet was used by governments on economic negotiations. Presently there is quite an amount of business that is conducted through Internet negotiation, and this figure keeps increasing (Mahadevan 1999). With the increasing adoption of Internet, it can be predicted that more and more business will be conducted through Internet. It is quite probable that in future Internet negotiation will be largely anonymous, which means that the organizations that the parties represent may be known but the negotiators' social, cultural and education backgrounds will not be known. This is especially probable in multicultural societies as Canada and the US.

The new technology facilitates the deployment and use of negotiation support tools. More advanced tools can provide a comprehensive and active support for conflict resolution based on decision and negotiation support systems (Guttman, Moukas et al.1998; Lo and

Kersten 1999; Sandholm 1999; Strobel 1999; Kersten and Noronha 2000). Models and support systems had already been launched to help negotiators reach mutually optimized results, such as the RAINS system used in the EU transboundary air pollution negotiations (IIASA 1998) and the simulation system used in the deep sea negotiation (Sebenius 1984; Nyhart and Goeltner 1987). Therefore, it is of value to study the behaviors and patterns of negotiation on the Internet. However, as a relatively new business activity, Internet negotiations have not been studied thoroughly.

Recently a project named InterNeg has been initiated in order to study the use of software tools in Internet negotiations, including cross-cultural and anonymous negotiations. The goal of this InterNeg project is to develop a comprehensive negotiation support system (NSS) on the World Wide Web. (http://interneg.org) Inspire is one of the systems developed within the InterNeg project (http://interneg.org/Inspire). The Inspire system allows the support of bilateral negotiations based on a mock business case. (The case presents a parts manufacturer positioned to supply components to a bicycle producer and will be discussed in more detail in later sections). This system offers the data to analyze and study Internet negotiation behaviors.

Reviewing the relevant literature reveals that there are not many studies of this social interaction and few hypotheses that explain the nature of Internet negotiations. Therefore we could not assume models or hypotheses from former literature. We adopted the approach of empirical modeling in this study. Tukey (1977) and Romesburg (1990) pointed out that empirical modeling prior to confirmatory analysis could aid the construction of significant hypotheses that may not have been apparent through confirmatory analyses alone. This indicates that it is advisable to establish new and meaningful models about Internet negotiations through empirical modeling methods, rather than constructing hypotheses from the literature and domain expertise alone. The recently developed data mining software packages offer us the tools to conduct the exploratory and empirical models. Data mining enables us to conduct exploratory analysis and detection of patterns (Chen, Han et al., 1996; Fayyad 1996). Therefore we made a study that is empirical and applied rather than theoretical. Different data mining methods were adopted to analyze the data.

The use of data mining is a recent approach to empirically model construction and validation. There are different methods and systems available that can be used to study the same data. Because it is not know if a particular method yields better (e.g., more expressive or robust) results, we decided to use several methods. In addition, using different methods may allow extracting richer information. Results obtained from using different methods may be the same or complementary.

This paper introduces the Inspire system, discusses the data collected from the web-base Inspire negotiation, the data mining methods used to study the data and presents results regarding both the data mining and the effectiveness of the methods. After applying 3 data mining methods with 16 different models on the data collected through Inspire system, some behavior patterns were found. These behaviors can be used to effectively increase the prediction accuracy of Internet negotiation results. We also made comparisons of different data mining models/methods' prediction accuracy.

In the initial study we found that the number of offers sent, and the time they had been sent has an impact on the achievement of an agreement. That gave researchers a deeper understanding of Internet negotiations. By applying three different data mining methods on real data and analyzing their prediction accuracies and time consumption, we found that some methods have higher prediction accuracy than others. While this result cannot be generalized, these methods may first to be used in similar studies in future.

In this study three data mining methods are used:

- Logistic Regression Analysis (LRA)
- Tree and Rule Induction (TRI)
- Artificial Neural Network (ANN)

These three techniques were chosen because: (1) the target variable of Internet negotiation is binary (agreement reached or not); and (2) many of the data set collected through the Inspire system are nominal or ordinal. The three selected techniques are often used on such kinds of data sets (Spangler, 1999). Furthermore, they have been widely researched and applied (James 1985; Limb and Meggs 1994; Weiss and Indurkhya 1998). An important factor is that many existing commercial software packages implement one or all of these techniques, including the packages that were available for this study. These commercial packages include SPSS Answer Tree, SAS Enterprise Miner, IBM Intelligence Seeker... (SPSS 1998; IBM 1999; SAS 1999).

Prior to this study, Mahadevan had conducted a study to analyze the Inspire data with different data mining techniques (Mahadevan 1999). The present study builds on

Mahadevan's work and extends it into two main directions. Firstly, different data mining techniques were studied; secondly, the focus is on the managerial assessment of data mining techniques rather than on statistical analysis.

The paper comprises six parts:

- Literature review introduces negotiation support and negotiations conducted over the Internet. The brief explanation of data mining and the three adopted data mining methods were also presented.
- The Inspire system is discussed in the second part, including a description of the system's usage, data collection, and the variables the data describers.
- Research Methodology provides detailed discussion of the data mining models and their application. The design of the research is also presented here.
- Implementation of data mining with SAS Enterprise Miner. This part gives brief description of how to conduct our experiment with the adopted data mining tool – SAS Enterprise Miner.
- Findings present the results of this study. This part consists of the Internet negotiation patterns found by different data mining methods, and the prediction accuracy comparisons of these three methods. Besides, the effect of layer number and hidden unit numbers in the layers on prediction accuracy and the time consumption of these methods are also analyzed and compared.
- Benefits and limitations discuss the benefit of this study to users and researchers, and this study's validity, reliability and limits. Some future potential researches are also presented.

2 Research Objectives

Internet negotiation is a new phenomenon hence, there are few researches conducted in this field and few models or hypothesis to explain it. This study is empirical and applied rather than theoretical. We used several data mining models in order to explore the data. Since data mining is still a new approach in extracting information (knowledge) from data, in addition to the theoretical comparisons of techniques, comparisons of results obtained by applying different methods is also useful. More than one method being used may allow extracting richer information. The overall goodof this paper is a better understanding of Internet negotiations supported by some specific tools and of the methods of data mining.

Specifically, the two objectives of this paper are:

1. To search for descriptive relationships and hidden patterns of Internet negotiation behavior by applying data mining tools.
2. To experimentally evaluate and compare the results obtained from different data mining methods from the managerial perspective, including predictive accuracy and time consumption.

3 Literature Review

3.1 Negotiation Support and Internet Negotiation

A broadly accepted definition of negotiation is that 'a negotiation is the process by which at least two parties try to reach an agreement on matters of mutual interest' (Hendon,

Hendon et al., 1996). Another definition is that of an operation where two or more parties with common (and conflicting) interests enter into a process of interaction with the goal of reaching an agreement (preferably of mutual benefit) (Lax and Sebenius 1986). The negotiation process is a process that allows the parties to redress their needs through mutual communications.

Recently people began to use negotiation support tools to increase the efficiency of negotiation, especially the computer-supported decision support systems (Jain and Solomon 2000). The research on the use of computing tools to conduct and support negotiations can be dated to the early 1980s (Chatterjee and Lilien 1984; Kersten 1985; Bui 1985). Lately, more and more negotiation support tools have been developed. Bui (Bui 1994) discussed some situations where computers could be used to support the face-to-face negotiations. Negotiation support systems (NSS) have also been used to develop efficient outcomes of negotiations to make the final agreements Pareto-optimal to all negotiating parties (Mahadevan 1999). Experiments on the use of NSS to improve the negotiations were conducted (Rangaswamy and Shell 1994). However, because it is a recently emerging business behavior, there are few theories explaining Internet negotiation.

Factors that can predict a negotiated agreement have been well studied in face-to-face situations (Adler, Brahm et al., 1992; Drake 1995). For example, Graham et al., measured bargaining strategies in ten cultures, in experiments of face-to-face negotiations, and concluded that opponent characteristics, such as personality, affect reaching negotiated agreements in certain cultures (Graham, Mintu et al. 1994). Since

Internet negotiations are a newly studied social phenomenon, there are few studies that explain the factors for reaching agreements in these negotiations. This condition makes it very difficult to assume the hypothesis. Therefore, using the empirical modeling method to study Internet negotiation became another choice, and the existing commercial data mining software enabled the feasibility of conducting it.

This thesis adopted exploratory data analysis rather than confirmatory analysis. It explored Internet negotiation patterns by applying 3 different data mining methods to the real data describing Inspire negotiators. These data mining techniques are typically used in large data sets to construct descriptive models (Chen, Han et al. 1996).

3.2 Data Mining

Data mining is a colloquial term used to address the empirical construction and verification of models from datasets containing a large number of elements (variables and observations) (Glymour, Madigan et al., 1997). The term "mining" is an artifact used to denote the aspect of analyzing data for information that may not have been extracted through the use of confirmatory analysis alone. Some data mining methods like TRI can construct models to meet intuitive explanation. Empirically verified models are assumed to be generalizable to the population outside the given data set (Mahadevan 1999).

Data mining is used to construct summaries of large amounts of data, to identify hidden structures and relationships, and to construct predictors of future observations (Harris-Jones 1997). Specifically, some tasks can be performed through using data mining techniques (Mahadevan 1999), including:

1. Predicting the class an observation belongs to;
2. Predicting the dependent variable value given independent variable values;
3. Formulating and describing clusters of similar observations;
4. Describing a group of observations;
5. Finding and describing relationships and associations among variables;
6. Identifying deviations and changes;
7. Identifying variables that control values of other variables.

Techniques used in data mining are typically non-parametric in nature. They were developed by the incorporation of statistical methods into machine learning (ML) algorithms (Limb and Meggs 1994). Machine learning research has also developed unique techniques for the purposes of data mining.

3.2.1 Data Mining and Knowledge Discovery in Databases (KDD)

Data mining is one important step in the process known as knowledge discovery in databases (KDD) (Fayyad, Piatetsky-Shapiro et al., 1996). Wirth, Shearer et al., (1997) proposed that knowledge discovery is used to extract new rules and patterns from a database. Fayyad (1996) observed that knowledge discovery in databases is a nontrivial process of identifying models that are valid, useful, and understandable by domain experts and users. He viewed the research of KDD as a multidisciplinary activity that encompasses techniques and issues beyond the scope of any one particular discipline. In general, KDD is used to identify significant patterns in data or to map low-level data into compact, abstract models.

The KDD process should include various stages of collecting raw data, process data - for use by applications - select and use data mining techniques, and evaluate constructed models. A typical KDD process includes the following seven steps (Adriaans and Zantinge 1996):

1. Initializing the problem and scope
2. Selecting Data
3. Cleaning/standardizing Data
4. Data enrichment
5. Coding
6. Data Mining
7. Reporting

KDD stages are iterative and interactive (John 1997). Although there is some canonical order to the process (e.g. the collection of raw data precedes other stages), information gathered at one stage may be used to alter assumptions at prior stages or to define further prior scopes (Mahadevan 1999).

3.2.2 Comparison of the Adopted Data Mining Methods

Three different data mining methods - tree and rule induction (referred to as TRI in flowing parts), logistic regression analysis (referred to as LRA in following parts) and artificial neural networks (referred to as ANN in following parts) - were adopted to analyze the data of our study. The reason why they were chosen was because that they are typical data mining methods used for binary or nominal target variables. Our research target was binary: Yes or No to represent if the agreement was reached, (in the data set

Yes and No were recorded as 1 and 0 separately). Therefore we chose these three methods. Furthermore, all these three methods were available in our software package, SAS 8.1 version.

Table 1 gives general ideas of the main characteristics of these three data mining methods. Usually a data mining technique can be categorized according to its method type, learning approach and statistical characteristics (Spangler, 1999).

Table 1. Comparison of Three Adopted Data Mining Methods

	Decision tree induction	Artificial neural networks	Logistic regression analysis
Type of method	Logic-based	Math-based	Math-based
Learning approach	Supervised	Supervised	Supervised
Linearity	Linear	Nonlinear	Nonlinear
Representational scheme	Set of decisions nodes and branches; production system	Functional relationship between attributes and classes	Functional relationship between attributes and classes

Each of the three methods is a popular data-mining method sharing a number of common characteristics while also exhibiting some notable differences (see Table 1). Weiss and Indurkhya divided data-mining algorithms into three groups: math-based methods, distance-based methods, and logic-based methods (Weiss and Indurkhya 1988). LRA is one of the most common math-based methods for nominal targets, as well as being the most common regression technique in general. TRI is the most common logic-based method. ANN is an increasingly popular nonlinear math-based method. Tools employing these methods are commonly available in commercial computer-based applications.

3.2.3 Supervised vs. Unsupervised Learning Techniques

In table 1, the term "Supervised" learning approach needs further explanation. All three methods adopted by this study are supervised learning techniques. That is, they induce rules for assigning observations to predefined classes from a set of examples, as opposed to unsupervised techniques, which both define classes and determine classification rules (Quinlan 1993; Shavlik and Dietterich 1990). Supervised learning techniques are appropriate to our study because the classes (results of negotiation) were defined exogenously and could not be modified by the decision-makers. Cluster analysis is an example of unsupervised learning algorithms. The methods also differ, particularly in the way they model the relationships among attributes and classes. Classification structures like LRA are expressed mathematically as a functional relationship between weighted attributes and resulting classes. TRI represents relationships as a set of decision nodes and branches, which, in turn, can be represented as a production system, or set of rules. LRA and ANN are nonlinear approaches while TRI is linear.

Liang argues that, because the choice of a learning method for an application is an important problem, research into the comparison of alternative (or perhaps even complementary) methods is likewise important (Liang 1992). That is especially true for data mining, where the costs and potential benefits involved strongly motivate the proper choice of methods in getting best results.

3.3 Data Mining vs. Confirmatory Statistical Methods

Data mining techniques have been contrasted with confirmatory statistical methods in the literature (Harris-Jones 1997; Pregibon 1998; Weiss and Indurkhya 1998). Confirmatory statistical methods allow for sophisticated analysis, however, they require a significant amount of end-user expertise. Further, they are also difficult to use in situations when the given data set contains very a large number of data elements (observations and variables) of which only a few may be significant to the required model. Thirdly, when the user cannot make assumption or hypothesis, data mining can help them to explore hidden patterns. The last two reasons are why we choose data mining in our research.

In comparison to confirmatory statistical methods, data mining techniques are used in the following situations (Mahadevan 1999):

1. **Unknown or complex data.** Data mining is used when the underlying distribution of data is unknown or complex (Glymour, Madigan et al., 1997). This allows for the derivation of significant information without knowledge of exact distributions.

2. **Sample is the population.** Data mining is used to make generalizations from a given set of data such that constructed models can be used to describe the population or predict the behavior of its members. While generalization is also a key issue in statistics, there is a significant difference between the two (Fayyad 1996; Mahadevan, Ponnudurai et al. 1999). In statistics, the issue of generalization involves the

definition of a population to which one can generalize and bring in issues of sampling. Without restrictions on the sampling method one cannot guarantee the statistical significance of the models. However, data mining ignores the population that is not described by a set of given data. There is an assumption here that population is in the data set, or sample is the population. Thus the results of data mining from a certain database may not apply to other ones.

3. **Different verification methods.** Model verification in data mining is conducted through sampling methods such as cross-validation and hold-out samples (Kohavi 1995). In these methods, observations from given data are used to verify the generalizable nature of constructed models, while statistical method usually verifies through new data.

3.4 Evaluating Data Mining Performance

Judging the performance of one data mining method over another requires the consideration of several modeling objectives. Spangler suggested that 3 aspects of each data mining technique should be evaluated from a managerial perspective (Spangler, 1999). They are prediction accuracy, comprehensibility, and speed of training and classification.

3.4.1 Predictive Accuracy

Most of the comparative studies we cited above measured the predictive accuracy and error rate of each method. Messier and Hansen, for example, compared the percentage of correct classifications produced by their induced rule system, to the percentage drawn from discriminant analysis, as well as individual and group judgments (Messier and Hansen 1988). It is difficult to make general claims about the relative predictive accuracy of the various methods. Performance is highly dependent on the domain and setting, the size and nature of the data set, the presence of noise and outliers in the data, and the validation technique(s) used. Predictive accuracy tends to be an important and prevalent indication of a method's performance.

3.4.2 Comprehensibility

Henery used this term to indicate the need for a classification method to provide clearly understood and justifiable decision support to a human manager or operator (Henery, Michie et al., 1994). TRI methods, because they explicitly structure the reasoning underlying the classification process, tend to have an inherent advantage over both traditional statistical classification models and ANN - which is almost a black box to users, since after the training process the user still are not able to understand the data patterns or know the effective predictors. Tessmer et al., argued that, while the traditional statistical methods provide efficient predictive accuracy, "*they do not provide an explicit description of the classification process.*" (Tessmer et al., 1993). Weiss and Kulikowski suggested that any explanation resident in mathematical inferencing techniques is buried in computations that are inaccessible to the "mathematically uninclined."(Weiss and Kulikowski 1991) The results of such techniques therefore might be misunderstood and misused. Rules and decision trees generated by TRI method, on the other hand, are more compatible with human reasoning and explanations.

3.4.3 Speed of Training and Classification

Speed can be an important consideration in some situations (Weiss and Kulikowski 1991). Henery suggested that a number of real-time applications, for example, must sacrifice some accuracy in order to classify and process items in a timely fashion (Henery, Michie et al., 1994). Again, because of situational dependencies, it is difficult to make generalizations about the computational expense of each method. For example, it is estimated that ANN methods that using back-propagation may require an unacceptably large amount of time (Weiss and Kulikowski 1991). (All the ANN models in SAS Enterprise Miner are using forward-propagation.)

Spangler et al. pointed out that not all methods are capable of estimation parameters for each of the representations (Spangler et al., 1999). Our strategy was to evaluate each method from a decision support perspective; especially the predictive accuracy and time consumption aspects, while conducting the case study on Internet negotiation with the Inspire system.

3.5 Tree and Rule Induction (TRI)

TRI (also named decision tree algorithm by some researchers) is attractive because its explicit representation of classification as a series of binary splits, makes the induced knowledge structure easy to understand and validate. TRI constructs a tree, and the tree can be translated into an equivalent set of rules.

We used SAS Enterprise Miner (referred as SAS EM in flowing parts) tools to apply the data mining methods. SAS software is currently widely regarded as one of the best and most powerful data analysis tool kits. Its data mining tool, named Enterprise Miner, enabled us to apply all 3 data mining methods. TRI induces a decision tree from a table of individual cases (usually saved in Excel format tables), each of which describes identified attributes. At each node, the algorithm builds the tree by assessing the conditional probabilities linking attributes and outcomes, and divides the subset of cases under consideration into two further subsets so as to minimize entropy...etc according to the criterion it chooses. The criterion for evaluating a splitting rule may be based on either a statistical significance test - namely an F test or a Chi-square test - or on the reduction in variance, entropy, or Gini impurity measure. The F test and Chi-square test accept a p-value input as a stopping rule. All criteria allow the creation of a sequence of sub-trees. The user may then use validation to select the best sub-tree. Each one has its benefits and weaknesses. The user can specify parameters that control the stopping behavior of the method before running the analysis.

If the training set contains no contradictory cases—that is, cases with identical attributes that are members of different classes—a fully-grown tree will produce an error rate of zero on the training set. Weiss and Kulikowski show that as a tree becomes more complex, measured by the number of decision nodes it contains, the danger of overfitting the data increases, and the predictive power of the tree declines commensurately (Weiss and Kulikowski 1991). That is, the true predictive error rate measured by the performance of the tree on test data becomes much higher than the error rate reflected in the performance of the tree against the training data alone. To minimize the true error rate, TRI models in SAS EM usually separate data to be mined into at least two parts:

training and validation, so as to minimize the predictive error by comparing the results from the training and validation data. Sometimes one more part-test data is also created to assess data mining models.

An empirical tree represents a segmentation of the data that is created by applying a series of simple rules. Each rule assigns an observation to a segment based on the value of one input. One rule is applied after another, resulting in a hierarchy of segments within segments. The hierarchy is called a tree, and each segment is called a node. The original segment contains the entire data set and is called the root node of the tree. A node with all its successors forms a branch of the node that created it; the final nodes are called leaves. For each leaf, a decision is made and applied to all observations in the leaf. The type of decision depends on the context. In predictive modeling, the decision is simply the predicted value.

TRI methods enable users to create decision trees that either: (SAS on-line documentation)

- Classify observations based on the values of nominal, binary, or ordinal targets,
- Predict possibilities of outcomes for interval targets, or
- Predict the appropriate decision when you specify decision alternatives.

An advantage of TRI methods over other data mining modeling methods, such as LRA and ANN, is that it produces a model that may represent interpretable English rules or logic statements. For example, "If monthly mortgage-to-income ratio is less than 25% and months posted late is less than 1 and salary is greater than $35,000, then issue a silver card."

Another advantage of the TRI method is the treatment of missing data. The search for a splitting rule uses the missing values of an input. Surrogate rules are available as backup when missing data prohibits the application of a splitting rule.

TRI methods produce a set of rules that can be used to generate predictions for a new data set. This information can then be used to drive business decisions. For example, in database marketing, TRI methods can be used to develop customer profiles that help marketers to target promotional mailings in order to generate a higher response rate.

The SAS implementation of TRI methods finds multi-way splits based on nominal, ordinal, and interval inputs. The user can choose the splitting criteria and other options that determine the method of tree construction. The options include the popular features of CHAID (Chi-squared automatic interaction detection), and those described by L. Breiman et al. in 1984. (Breiman et al., 1984)

The TRI method in SAS EM (SAS on-line documentation):

"...supports both automatic and interactive training. When you run this method in automatic mode, it automatically ranks the input variables based on the strength of their contribution to the tree. This ranking may be used to select variables for use in subsequent modeling. In addition, dummy variables that represent important "interactions" between variables can be automatically generated for use in subsequent modeling. You may override any automatic step with the option to interactively define a splitting rule and prune explicit nodes or subtrees. Interactive training enables you to explore and evaluate a large set of trees that you develop 'on the fly' ".

3.6 Artificial Neural Networks (ANN)

Artificial neural networks simulate human cognition by modeling the inherent parallelism of neural circuits found in the brain using mathematical models of how the circuits function (Spangler et al., 1999). ANN methods were originally developed by researchers who tried to mimic the neurophysiology of the human brain. By combining many simple computing elements (neurons or units) into a highly interconnected system, these researchers hoped to produce complex phenomena such as intelligence. In recent years, neural network researchers have incorporated methods from statistics and numerical analysis into their networks. While there is considerable controversy over whether artificial neural networks are really intelligent, there is no doubt that they have developed into very useful statistical models (SAS on-line documentation). More specifically, feedforward neural networks are a class of flexible nonlinear regression, discriminant, and data reduction models. By detecting complex nonlinear relationships in data, neural networks can help to make predictions about real-world problems.

Neural networks are especially useful for prediction problems where (SAS On-line Documentation):

- No mathematical formula is known that relates inputs to outputs.
- Prediction is more important than explanation.
- There is lots of training data.

Common applications of neural networks include credit risk assessment, direct marketing, and sales prediction. Up until the early 1990s, neural networks were often viewed as alternatives to statistical methods. Some researchers made claims that neural networks could be used to analyze data with no expertise required on the part of the analyst. These unjustifiable claims, combined with the unreliability of early algorithms such as standard backprop, led to a backlash in which many people, especially statisticians, dismissed neural networks as being worthless for data analysis. But in recent years, it has been widely recognized that many kinds of neural networks are statistical methods, and that when neural networks are trained via reliable methods such as conventional optimization techniques or Bayesian learning, the results are just as valid as those obtained by many nonlinear or nonparametric statistical methods.

Neural networks, like other statistical methods, cannot magically create information out of nothing - the rule "garbage in, garbage out" still applies. The predictive ability of a neural network depends in part on the quality of the training data. It is also important for the analyst to have some knowledge of the subject matter, especially for selecting inputs and choosing an appropriate error function. Experienced neural network users typically try several architectures to determine the best network for a specific data set. The design process and the training process are both iterative.

A neural network consists of units (neurons) and connections between those units. There are three kinds of units: (SAS On-line Documentation)

1. Input units: Obtain the values of input variables and optionally standardize those values.

2. Hidden units: Perform internal computations, providing the nonlinearity that makes neural networks powerful.

3. Output units: Compute predicted values and compare those predicted values with the values of the target variables.

All the units in a given layer share certain characteristics in SAS EM. For example, all the input units in a given layer have the same measurement level and the same method of standardization. All the units in a given hidden layer have the same combination function and the same activation function. All the units in a given output layer have the same combination function, activation function, and error function. A network may contain many units. In SAS EM, the units are grouped into layers to make them easier to manage. There can be several input layers, several hidden layers, and several output layers. In the SAS EM Neural Network models, when you connect two layers, every unit in the first layer is connected to every unit in the second layer.

An ANN model is specified by defining the number of layers it has, the number of nodes in each layer, the way in which the nodes are connected, and the nonlinear function used to compute node values. Estimation of the specified model involves determining the best set of weights for the arcs and threshold values for the nodes.

An ANN is trained - that is, its parameters are estimated - using nonlinear optimization. The network propagates inputs through the network, derives a set of output values, compares the computed output to the provided (corresponding) output, and calculates the difference between the two numbers (i.e., the error).

If there exists a difference, the algorithm proceeds backward through the hidden layer(s) to the input layer, adjusting the weights between connections based on their gradients to reduce the sum of squared output errors. When the total error is acceptably small, the algorithm stops.

Most connections in a network have an associated numeric value called a weight or parameter estimate. The training methods attempt to minimize the error function by iteratively adjusting the values of the weights. Most units also have one or two associated numeric values called bias and altitude, which are also estimated parameters adjusted by training methods.

While inputs to an ANN might be integer or discrete, the *"weighted nonlinear transformations of the inputs as part of their being fed forward through the network result in continuous output level values. ... Continuous output levels result in a more tractable error measure for the back-propagation algorithm to optimize and also permit the interpretation of outputs as partial group membership. Partial group membership means an ANN is capable of representing inexact matching, if that is the way to find a best fit for some set of input data. It also can model classification tasks that are inherently 'fuzzy'—that is, tasks that generally are simple for humans but traditionally difficult for computers.*" (Help File of SAS Version 8.1)

ANN may be difficult to specify because of their flexibility. Adding too much structure to an ANN makes it prone to over-fitting, but too little structure might prevent it from capturing the patterns in the data set. Our study will check if its structure affects the results. Those patterns are represented in the connection weights and the node thresholds; a form that is not transparent to users. Computationally, if the training set is large, back-propagation and related algorithms may require a lot of time. Neural networks are often treated as a black box, with only the inputs and outputs visible to the decision-maker. The classification chosen by the ANN is not easily visible to the user, and the decision process that led to that classification is not. That's why the ANN process is often called a "black-box" process.

The most popular form of neural network architecture is the multi-layer perceptron (MLP), which is the default architecture in SAS EM and was the architecture we applied. A multi-layer perceptron (SAS on-line documentation, version 8.1):

Has any number of inputs;

- Has one or more hidden layers with any number of units;
- Uses linear combination functions in the hidden and output layers;
- Uses sigmoid activation functions in the hidden layers;
- Has any number of outputs with any activation function;
- Has connections between the input layer and the first hidden layer, between the hidden layers, and between the last hidden layer and the output layer.

SAS claims that "given enough data, enough hidden units, and enough training time, an MLP with just one hidden layer can learn to approximate virtually any function to any degree of accuracy" and "for this reason MLPs are known as universal approximators and can be used when you have little prior knowledge of the relationship between inputs and targets". (Help File of SAS Version 8.1).

3.7 Logistic Regression Analysis (LRA)

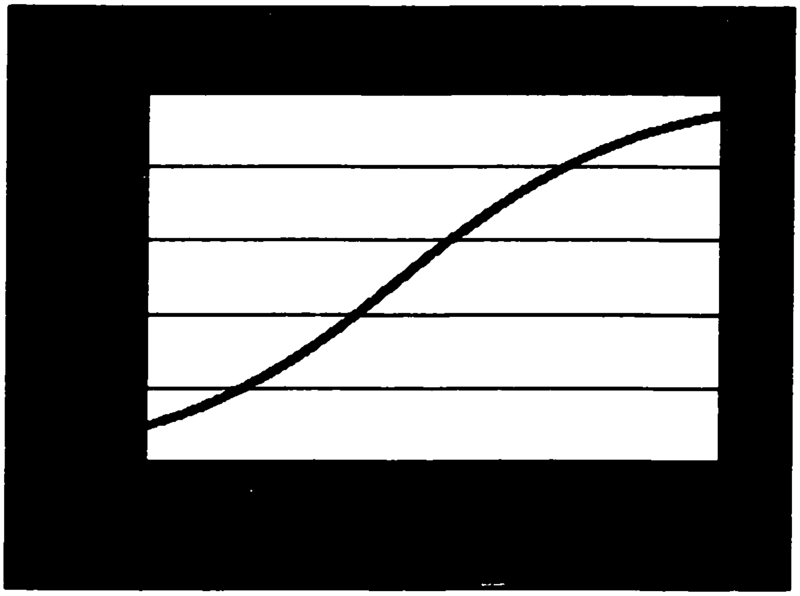
Regression has been the most popular statistical technique in reality; many data mining software also include it. Regression can be classified into two general types: linear and nonlinear. Logistic regression belongs to nonlinear regression and is used while the dependent variable is discrete. It can be categorized as simple logistic regression if the dependent variable is binary, or multiple logistic regression if the dependent variable is categorical (nominal)

Logistic regression estimates the probability of a certain event, such as a loan default, occurring. It uses observed factors coupled with occurrences or non-occurrences of the event to model the probability of occurrence under different factor conditions (SPSS, 2000).

Following is an example and diagram supplied by SPSS to explain the logistic regression technique used in data mining (http://www.spss.com/datamine/logistic.htm).

Suppose we believe that a loan amount has an effect on the probability of a default. A logistic regression can tell whether or not there is a significant effect from the loan amount on the loan default. If a significant effect is found, the bank can use this in a model estimating the probability of defaults on new loans and then decide whether or not to grant the loan. The figure below shows the fitted probabilities as a function of the loan amount.

Since the most interesting dependent variable in our Internet negotiation data set is binary (only 2 discrete statuses are valid: agreement or no agreement), we decided to use the logistic regression to find effective predictors.



(Source: SPSS web site: http://www.spss.com/datamine/logistic.htm, November 2000)

4 The Inspire System and Data Description

Advancements in network technology - such as the Internet and World Wide Web - have led to the development of net-centric tools (Carroll 1996). Such tools, like Microsoft NetMeeting...etc are now facilitating new forms of communications, including the possibility of communicating and negotiating through the Internet.

The InterNeg project (http://interneg.org) was initially developed in 1995 at the Center for Computer Assisted Management (CCAM) and now is hosted at Concordia and Carleton Universities. The objectives of the project are to *"Design and implement negotiation support system systems which integrate support aids with easy and rapid communication facilities. ... Establish an international network of collaborators to help develop, test and deploy negotiation support systems."* ...etc., which includes the study of Internet negotiations. Through the InterNeg project, researchers develop and host a comprehensive negotiation support system (NSS) on the World Wide Web. Inspire, one of the negotiation support systems developed by the InterNeg group, provides support for the negotiation process through electronic bargaining facilities, visualization suites, analytical tools, and quantitative and qualitative models (Kersten and Noronha 1999).

Inspire was developed to meet both training and research requirements. For research purposes, tools were developed to capture data at various points in the negotiation process that is available to researchers in study aspects of Internet negotiations.

We preferred using the real data set collected through the Inspire system to the use of generated data in evaluating the data mining techniques. This is because it is representative of an important class of business problems, and because the noise in the data set provides us with information on the sensitivity of the approaches to "dirt" in a data set. Artificial data would have allowed us to carefully control the population parameters from which the data are drawn, but would have required that we first define and estimate values for all such critical population parameters. Numerical research can be done with artificially generated data or with real data. Conclusions based on empirical research are useful if the characteristics of the samples on which they are based are sufficiently similar to those of problem instances others are likely to encounter in practice (Spangler et. al., 1999).

4.1 The Negotiation Process with the Inspire System

A business negotiation case involving a seller and a buyer is conducted in Inspire. This case is a mock scenario and it is the same for every negotiation. One side of the negotiation represents a parts supplier (called Itex Manufacturing) positioned to supply key components to a bicycle producer (called Cypress Cycles), Participants are assigned the role of bargainer for either side and negotiate on behalf of the company. The Inspire system is used by the participants to communicate so as to reach an agreement that would maximize their party's position. Anonymity between the participants is further strengthened since they are not made aware of their opponent's identity a priori (although they may elect to reveal their identity during negotiations).

Inspire negotiations include three phases: the analysis, the conduct of negotiation, and the post-settlement phase. The stages are reflective of tasks involved in negotiation studies involving face-to-face negotiations.

4.1.1 The Analysis Phase

The analysis phase involves an analysis of the situation, the problem and the opponent, formulation of preferences, reservation levels, and strategy. As part of this phase, the user specifies their preferences over four distinct terms. (For more details refer to http://www.interneg.org/inspire or Kersten and Noronha 1999).

1. **Price.** The unit price that Cypress would pay to buy the parts from Itex. The user can select one of the five options: $ 3.47, $ 3.71, $ 3.98, $ 4.12 or $ 4.37.

2. **Delivery**. The delivery time of the parts from Itex to Cypress. The user can select one of four options: 20 days, 30 days, 45 days or 60 days.

3. **Payment.** The payment schedule for payment of the parts. The user can select one of three options: payment on delivery, payment 30 days after delivery or payment 60 days after delivery.

4. **Return.** The terms for the return of defective parts. The user can select one of three options: full price returns, 75% refund with 5% spoilage or 75% refund with 10% spoilage.

The negotiator needs to rate each term and, subsequently, each option available under an item on a 100-point scale according to his or her preference. The option ratings are, typically, dependent on the side the user represents—for example, the seller may put more emphasis on a higher price, slower delivery, shorter payback, and strict terms for returning defective parts. Then a utility function is built by the Inspire system using the user's ratings. The users then review the derived ratings and may modify them.

4.1.2 The Negotiation Phase

The negotiations phase in Inspire negotiations involves the exchange of offers, messages, and offers with messages by users. Offers always include the negotiated items and their options, such that the participant provides one option per item. The Inspire system offers negotiation support during this phase by automatically presenting ratings for each offer and counter-offers based on the utility function which was constructed based on the user's preferences. The user, therefore, can more easily make decision while using these ratings to construct offers and to evaluate counter-offers from their counterparts.

4.1.3 Post-settlement

The post-settlement phase only involves the evaluation of the negotiation outcomes generated by, and after, negotiation activity. These outcomes include information about the agreement and the negotiators' satisfaction. Furthermore, Inspire users have the possibility of improving inefficient compromises because the system will automatically check their utility table. If there is a possible option that can make at least one side more satisfied while not damaging the other, the system will offer this new option to help them reach the optimal choice for both sides.

The negotiation finishes either when one of the three following events takes place: (1) two users reach an agreement; (2) when a negotiator terminates the mediation; or (3) when the deadline has been met.

4.2 Data Collection in the Inspire System

The data collected from the Inspire system is unique and describes the first experiment involving a large number of bilateral business negotiations carried out on the web. The Inspire system provides two sources of data to describe the entire negotiation and the users together:

1. Data collected by two questionnaires (pre and post negotiation) which are filled on-line by each negotiator

2. Data generated by the computer system to record the activities of the negotiation.

The first questionnaire is presented in the analysis phase of the negotiation. This is referred to as the "pre-negotiation questionnaire". The second questionnaire is presented in the post-settlement phase, and is referred to as the "post-negotiation questionnaire". The pre-negotiation questionnaire contains 16 questions about the user's background, such as the user's date of birth, their country of birth and residence, countries they have lived in, their mother tongue, their prior level of negotiation experience, etc.

The post-negotiation questionnaire contains questions about the system and the just-reached agreement, the questions asked about the user's opinion of the system, the negotiation process, and their counterparts, etc.

The history recording mechanism in the Inspire system describes each negotiator's activities in detail and provides complete computer records of the negotiation. This includes the user's preference ratings, the interaction measures between the two sides (for example, number of offers and messages exchanged), the length of the negotiations, and the score of the agreement (obtained by applying the negotiator's utility function to the final package)...etc.

4.3 Structure of the Data Collected through the Inspire System

The variables that describe the Inspire negotiations can be classified into three categories: exogenous variables, response variables, and intermediate variables (Kersten and Noronha, 1999). (Please refer to Appendix V to see the brief definitions of variables.)

4.3.1 Exogenous Variables

The first type of variable is the exogenous variable and contains 3 sub-categories: subject, task, and system variables.

Subject variables describe information about the subjects in this research. This includes measurements of country of birth, age, gender, occupation, and the prior negotiation experience of the user, etc.

Task variables relate to the problems undertaken by the subjects; these center around the negotiation case. Cases from different problem domains, or with different agendas have the potential for different effects on outcomes. For the convenience of research, all the users are given only the same case in Inspire, and thus this variable is held constant throughout.

System variables represent the specific functionality that are available from Inspire. This contains the ratings that appear when the user is constructing an offer, the negotiation history graph that displays the flow of past offers, and the item and option ratings constructed by the user, etc.

4.3.2 Response Variables

The second type is the response variable and it contains 3 sub-categories too: measures of goodness of the negotiation's outcome, measures of goodness of the negotiation process, and measures of the system's effectiveness.

Measures of goodness of the negotiation's outcome indicate the effectiveness of the negotiation. This can be assessed by the presence of an agreement, and whether this agreement was efficient.

Measures of goodness of the negotiation process are subjective to each negotiator. Therefore, the same negotiation process can have polar impacts on the two users.

Measures of the system's effectiveness are elicited through the post-questionnaire, where users are asked their perception of various system functionality, including the negotiation history graph, the score display for offers, and the in-built messaging systems (Mahadevan 1999).

4.3.3 Intermediate Variables

The third type of variable is the intermediate variable. It contains measures that have a contributory or intervening effect on the relationship between independent and dependent variables (Emory and Cooper 1991). Intermediate variables also contain three sub-categories: "psychological baggage", the behavior of negotiators during negotiations, and the perceptions of the negotiations.

The "psychological baggage" that a user brings to the table is dependent on the user's personal background and impacts the process and the outcome of negotiations (Kersten and Noronha 1999). For example, a user from a culture where competitive bargaining is normal may expect greater hostility from counterparts than users from a culture where relationship building is the basis for negotiation.

Users' behaviors during negotiations may have a direct impact on the negotiation process and outcome. For example, exchanging a low number of offers and messages may not lead to agreement between the users.

Perceptions of the negotiations may be affected by both the person's background and by the process and outcome of negotiations. Perceptions of actual control during negotiations, satisfaction with the final agreement, and perceptions of negotiation performance...etc are in this group of factors.

5 Research Methodology

5.1 Empirical Exploration of Internet Negotiation Behavior Patterns

As we discussed at the beginning, the conduct of anonymous negotiations through the World Wide Web (WWW) is a newly emerging social phenomenon that has not yet been well studied. Only a few hypotheses and studies on the nature of Internet negotiations and the interactions between underlying concepts and constructs have been reported in the literature (Kersten and Noronha 1999; Mahadevan, 1999). Data collected of Inspire negotiations offer the possibility of empirically modeling behavior relations and patterns of Internet negotiations.

The data set is a census of 1525 cases of Internet negotiations and containing descriptive information of 3050 participants. The data was collected by the Inspire system at web site http://interneg.org/Inspire from September 1995 to November 2000. Each negotiation case is recorded in two separate records, one for each negotiator. Each record contains 82 attributes describing one participant's demographics and actions.

This study attempts to determine significant predictors of the result of Internet negotiations. Since there are few hypotheses and literature on this topic, the data is explored with the data mining tools rather than confirmatory statistical analysis, which tests the hypotheses constructed from literature domain. It continues an earlier study reported by Mahadevan (1999), with the following differences in the treatment of data, and use and assessment of the tools: (1) The primary concern of Mahadevan's study was the comparison of the statistical models used in four data mining tools, including association, cluster, classification and decision tree. The objective of this study is first to determine the patterns or rules describing negotiators, and our data mining methods are logistic regression, artificial neural networks and decision trees. (2) The perspective taken in this study is more of a researcher in business or social science than a statistician or a computer scientist. We select methods that are different in order to compare their explanation and prediction powers rather than compare method application and difference in data treatment. We concern the managerial aspects of data mining including the prediction accuracy and time consumption, while he "compared differences and similarities present in models of anonymous negotiations constructed with association rules, naïve Bayes, entropy-based decision tree, CHAID techniques." (3) We have one specific target variable – the agreement reached or not while Mahadevan concerned varied dependent variables selected from the set of intermediate or response variables. (Mahadevan, 1999)

We first conduct three data mining methods on the Inspire data, and use their results to generate behavior patterns or rules of Internet negotiations. After that, we use the results of each data mining model to predict the known data cases. By comparing the predicted results with the actual results we can obtain the prediction accuracy of each data mining method/model. We repeat this by five times, which will give us enough data to conduct the mean comparison analysis to compare their accuracy. Our experiment model, in statistics, is called the single randomized block design. We then adopt the single factor with blocks ANOVA method to analyze that (referred as ANOVA here). While applying the mining process, we count the consumed time. The ANOVA analysis can also reveal if their time consumption is the same.

5.2 Data Mining Process and Models

In this study the typical data mining approach (SAS online documentation, Version 8.1) is used to find out Internet negotiation patterns. It includes following 9 steps:

1. Setting Target Profile
2. Sampling
3. Data Partitioning
4. Selection of Variables
5. Transformation of Variables
6. Setting Data Mining Models
7. Mining the Training Data
8. Model Assessment
9. Prediction

By applying the first 7 steps we are able to search hidden patterns of Internet negotiations. We need to apply the 9^{th} step to compare the prediction accuracy.

5.2.1 Setting Target Profile

The first step of data mining is to define the target variable and determine the target profile. In our study the target variable is the negotiation result (labeled as status of agreement in the data sets). Its value must be either of the 2 values: 0 or 1, where 0 represents no agreement reached and 1 represents that there is an agreement reached.

Since we only concern the prediction accuracy, the correct prediction of an agreement can be reached has the same weight as the correct prediction of an agreement cannot be reached. We define the prediction correctness to be 1 for each correct prediction and 0 for each incorrect prediction. Then the prediction correctness matrix is as follows:

Table 2. Profit Matrix of Prediction on Target

	Actual Value is 1	Actual Value is 0
Predicted Value is 1	1	0
Predicted Value is 0	0	1

While implementing the data mining methods, we make each method take a criterion to maximize the prediction correctness. Then each method will obtain some rules or predictors so that the prediction correctness can be maximized.

5.2.2 Sampling

In many cases, data mining deals with databases that contain millions of records. Therefore, a sampling process prior to conducting the "mining" process happens usually. In this step the number of cases selected and the sampling methods (including simple random, stratified…etc.) will be determined. After sampling the data for data mining is much less and the time consumed can be greatly reduced. Most data mining software requires a minimum amount of cases to conduct effective mining results, for example, SAS Enterprise Miner Version 4.1 (referred here as SAS EM) need at least 2000 cases to implement data mining effectively. Since we only have 3050 data (1525 cases), this step is not necessary for us. We apply data mining on the whole 3050 data.

5.2.3 Data Partition

After the sampling step, the selected data (referred as selected data in follows) need to be partitioned into 3 subsets. Both the first and the second subset are used for training and validating the results obtained from the other subset. Each data mining model mine both of them at the same time. The final result for output is the one that minimizes the sum of their errors. The third one is a test subset, which is used to assess the model's efficiency in step 8. The using of two interactive training and validating subsets increases the results' reliability than using only one training subset. The third subset, test subset, is optional and is used for evaluating the rules or predictors generated from the training and validation subsets. It offers an assessment of the result. In our study we implement the prediction step on known data to test the prediction accuracy, so the 8^{th} step – model assessment – is not necessary and skipped.

The user defines the percentage each one has in the selected data set. We partition the population into 40% (1220 cases), 40% (1220 cases), and 20% (610 cases) as first, second and test subsets respectively. We take a large proportion of data (80%) in the first and second subsets so as to obtain more reliable results. We randomly partition the selected data five times and therefore can obtain 5 different partitioning of data. Each data set is different from another, but all contains the training, validating and test subsets with respectively 1220, 1220 and 610 cases. Therefore the comparison of their results is more reliable than suing different size data sets. On each data set we apply all the 3 adopted data mining methods and record the results. It is understandable that the results may vary since each time we mine in different training and validation subsets since the training data are different. The results may be the same or complementary due to the differences between training and validation subsets.

5.2.4 Variable Selection

In this step different methods can be used to select and remove unwanted variables. Usually two factors need to be analyzed here. The first one is if the variables have enough values to conduct data mining. It means if their missing values exceeding an acceptable level. In our study we reject the variables with more than 40% of missing values (has less than 1830 valid records), because we need approximately 2000 cases to implement data mining effectively (SAS online documentation, version8.1). The second factor is if the variables are unrelated to the target variables. Two selection methods have been proposed (SAS online documentation, Version8.1): the R-square test and Chi-square test. After conducting the R-square and Chi-square methods, we find all the independent variables unrelated to the target variable by traditional statistical methods. If we apply to remove the variables according to this factor, then all variables would have been removed. Therefore we do not apply this rule (this is normal in data mining since this tool is mainly used to find out hidden patterns that traditional statistical methods can hardly get) but kept all the independent variables with missing values less than 40%.

Rather the second factor if a variable is obviously correlated to the target and the relation is not useful, we just remove it at this step, e.g. in a case if no offer was sent, certainly no agreement could be reached. This is obvious but useless.

5.2.5 Transformation of Variables

The original variables may need some transformations to increase the prediction power and the model's fitness. For example, the regression methods assume the variables are normally distributed. If the variable for mining is very skewed (not evenly distributed around the mean), then the results from regression methods are not reliable. Therefore reducing the skewness usually is the first variable transformation action. We found that most of the interval independent variables are skewed (refer to Appendix VI) and they would cause incorrect prediction. Therefore we make the "Maximizing Normality" transformation on almost all interval independent variables to make them approximately normally distributed (refer to Appendix VI to see how each variable is transformed, SAS EM can choose the proper transformation by itself.), except for the "Age" variable. Instead we divide it into four age intervals: under 20, 20-30, 30-40 and over 40 so as to more explicitly see if age difference can affect our target variable.

5.2.6 Setting Data Mining Models

In order to obtain more reliable results, we implement all the data mining methods to 5 different training and validation data sets. In step 3 we have obtained 5 sets of training and validation data sets (the first and the second subsets, from now on referred as training data). Each round all the data mining methods are applied on the same training data. We repeat the mine process five times till all 5 training data are mined. Therefore totally we have 5 rounds of data mining processes on 5 different training data.

Each round one training data is mined with 3 data mining methods (each data mining method contains more than one model). Both the results and the run time for each model are recorded and repeated. Then in next round another training data is used and mined. This process is repeated till the 5th round and all 5 training data are used. All the results and the run time are recorded so as to make comparisons.

Each data mining method contains different approaches inside it, or consists of many different structures (for ANN methods). We specifically design and implement 16 data mining models from the three data mining methods. Following sections give more detail on how the 16 models of 3 different data mining methods in this study are defined.

- **TRI:** There are many different algorithms used in TRI mining method: Chi-square test, Exhausted Chi-square test, CART, Entropy reduction and Gini reduction. We select 3 different TRI models in our study (since SAS EM 4.1 only offers these three); each operates a different algorithm, including the Chi-square, Entropy reduction and Gini reduction.

- **LRA:** There are three different ways to remove invalid predictors in LRA method: stepwise, backward and forward. Three different LRA models were used with the forward, backward and stepwise approach representatively.
- **ANN:** Two important factors, as discussed before, are important in ANN method: the number of hidden layers and the number of hidden units in each layer. These two factors can make many combinations. On the other hand, since ANN methods analyze the data in the black box, they do not give users understandable rules or predictors. The user can only use them to predict unknown data without knowing what they have learnt from the training data or how they make predictions.

ANN method does not offer explanations to users. Thus we cannot use this method to search the behavior patterns of Internet negotiation, but can only test their prediction accuracies. However, we are also interested in if the layer number and hidden unit number has impact on the prediction accuracy. Since when a user need to use ANN method to mine his data and make predictions, he has to set how many layers his models contains, how many hidden unit each layer contains and how the layers are structured. ANN models are mainly used to make predictions, so the users concern on how to maximize the prediction accuracy the most. Do different ANN models have different prediction accuracy? If so, which one is better? And what is the impact of the layer number and hidden unit number on the prediction accuracy?

In order to study this we design two groups of ANN models and each contains five different ANN models. In most cases the users of ANN methods do not have more than 5 layers or hidden units in their models (SAS online documentation, Version 8.1).

Therefore five different models to test their impacts are sufficient. In the first group we make each model contain different layer number, from 1 to 5 representatively, and set 3 hidden units in all the layers (since most users set 2 or 3 hidden units in their models). The second group also contains 5 ANN models. All of them have 2 layers inside, (We make all models contain 2 layers because our computer encountered many times of crash while running big ANN model with more than 3 layers and 3 hidden units), and each model has different hidden unit number in its two layers, from 1 to 5 representatively. The first group is designed to test the effect of layer numbers on prediction accuracy and the second group to check the effect of hidden unit number in the layers. (Please refers to Appendix VII to see how the two groups of ANN models are designed.)

Therefore we have 16 models in total. They are:

1. LRA stepwise
2. LRA backward
3. LRA forward
4. TRI Chi-square Test
5. TRI entropy reduction
6. TRI Gini reduction
7. ANN 3 hidden units, 1 layer (Group 1)
8. ANN 3 hidden units, 2 layers (Group 1)
9. ANN 3 hidden units, 3 layers (Group 1)
10. ANN 3 hidden units, 4 layers (Group 1)
11. ANN 3 hidden units, 5 layers (Group 1)

12. ANN 2 layers, 1 hidden unit (Group 2)

13. ANN 2 layers, 2 hidden units (Group 2)

14. ANN 2 layers, 3 hidden units (Group 2)

15. ANN 2 layers, 4 hidden units (Group 2)

16. ANN 2 layers, 5 hidden units (Group 2)

5.2.7 Mining the Training Data

In total we have 16 different data mining models: three from TRI, three from LRA and ten from ANN in two groups. We also have five equal size but different training data sets. Each round we take one training data set and apply all sixteen models on it. Each round all the sixteen models mine the same data. Thus the comparisons of mining results are valid. We record the time consumption that each model mines in the training data as well as its mining result. All the mining results from TRI and LRA methods are analyzed to extract effective Internet negotiation behavior patterns. In order to compare the prediction accuracy one more step – the predicting - has to be implemented. It is discussed in next section.

5.2.8 Model Assessment

After the mining process an assessment of the model's efficiency is necessary. The correct assessment is essential for the users to know how effective the model can make predictions. Only then can they make correct decisions. The test subset created in step 3 is for this purpose. However, since we make predictions on known data and we will know (and compare) the prediction accuracy, this step is not necessary in our study.

5.2.9 Prediction

In order to evaluate the prediction accuracy of each method, we need to make predictions using the data mining results. Usually the prediction step of data mining is used to predict unknown target by applying the rules found from a specific model to the known independent variable (SAS online documentation, version 8.1). In our study, instead of using the data that we do not know the negotiation results, we randomly select five different data sets from the population and each has 1221 cases (about 40%). These five data sets are used to make predictions on (they are referred as predicted data in follows), and we have known their negotiation results.

5.3 Compare Prediction Accuracy of Different Data Mining Methods

The most practical application of data mining tools is to make correct predictions. The increased prediction accuracy achieved by data mining certainly brings significant benefit to users. Thus it is reasonable that the prediction accuracy is one of the top concerns of users. Therefore while using data mining to search the behavior patterns in Internet negotiation, we are interested in the comparisons of their prediction accuracy.

All the sixteen models' results are used to make predictions on the predicted data. Following rules are obeyed in our study:

- All the training data are randomly chosen from the population and have the same size;
- All the predicted data are randomly chosen from the population and have the

same size;

- All the results obtained from one training data will be used to predict a same predicted data;
- Each predicted data can only be predicted by using the results obtained from a same training data.

Each round one training data is taken and we can obtain sixteen results from the sixteen different models. After applying them to predict a predicted data, we can also get sixteen prediction results in each round. We compare the predict results with the actual results, and record the prediction accuracy of each model in each round.

By alternating the results from each round and the predicted data, we repeat this process by five times till all data are used and record all the prediction accuracy. From each model we can obtain five prediction accuracy. Then we can use the ANOVA mean-compare method to check if there are significant differences between their prediction accuracy.

5.4 Compare the Processing Time of Different Methods

While repeatedly implementing every data mining method on the same data, we record the running time of each model's data mining step. Only the time spent on the data mining step is counted, other steps like sampling or partitioning are excluded since we are interested in the time consuming of different mining methods only. Because (1) each round all methods/models were mining in the same data set, and (2) the training data in each round are random chosen and in same size, thus the comparison of their time consumption is valid. We can get five time consumption data for each model from the five rounds. Then the ANOVA analysis can then reveal if there is any difference between their running time on mining process.

5.5 Hidden Units and Layers' Effect on Prediction Accuracy in ANN Method

The structure of an ANN model is defined by its hidden layer number and the hidden unit number inside the layers. Does the number of hidden layers and the hidden units in the layers affect the prediction accuracy? We design two separate groups of ANN models to find out the answer. In the first group we keep the number of hidden units in the hidden layers constant and the number of hidden layer varied from one to five. In the second group the number of hidden layers is constant and the number of hidden units in the hidden layers varied from one to five. From the five rounds of data mining processes we can obtain 5 prediction accuracy for each model. The ANOVA analysis on the first group models' prediction accuracy can reveal the layer number's effect; and analysis on the second group can show the hidden unit number's effect on prediction accuracy.

6 Implementing Data Mining with SAS EM

The commercial software, SAS Enterprise Miner (referred here as SAS EM), is chosen to apply the data mining methods on our Inspire data. SAS EM has been selected because (1) it is a well-known and popular software package; (2) it contains all the three methods we intend to use to conduct both the exploratory and empirical analyses.

The availability of the SAS EM and its wide use in business and other studies was also an important criterion for our selection of this software. It includes more data mining methods than other available software. For example, SPSS provides two programs Answer Tree and Clementine. The SPSS Answer Tree only offers TRI method and SPSS Clementine covers other including ANN and Cluster...etc., but not TRI methods. Since we are also interested in the factor that how time consuming each data mining methods is, having one software containing all these 3 methods may give us more reliable information because one may assume that the methods have been implemented in a similar manner. Lastly using one comprehensive software instead of several makes the conduct of experiment easy.

6.1 Input Inspire Data

Before implement the data mining steps, we need to input our Inspire data into the SAS libraries. SAS EM is a component of SAS 8.1 version. All the components have shared SAS data libraries. SAS EM works on and only on the internal data in SAS libraries. Since our Inspire data is saved in an external SPSS format file, it has to be re-saved as Excel 97 format and then imported into SAS library. In order to keep the data for next time use, users have to create his or her own library in SAS and make it to be "start-up" type by checking the check box beside, and then import the data into it. Otherwise when SAS program closes, it automatically cleans all the data out of its memory. Users cannot retrieve the data again next time.

After the data has been input, SAS EM shows the amount of cases in the data and the variable's attributes and separates all variables into two groups: interval and class variables. All the information here is used by users to decide how to select and transform variable.

6.2 Setting Target Profile in SAS EM

In its data input function SAS EM enables users to define the target Profile. After setting the target variable, then right click it, a profile window will comes out. We then define the profit matrix as above discussed and select the "Maximizing Profit" to be the mining criterion. As we have mentioned, in our case it equals to maximizing the prediction accuracy.

6.3 Data Partitioning, Variables Selection and Transformations in SAS EM

Since we have decided to mine the whole population, the step of sampling is not conducted in our study. SAS EM offers us the function to partition data. We simply indicate the desired percentages of training, validation and test subset in its dialog box, then SAS EM can do it automatically. In this step the user also needs to specify a seed number, which is used to decide how to randomly partition the selected data set. With different seed SAS EM can generate different subsets. Therefore we just created 5 different seeds in each round, SAS EM gave us different training, validation and test subsets each round. We kept the percentages unchanged, so all the subsets in different round have same sizes.

Similarly, we only need to input the desired level - less than 40% of missing values - into the variable selection dialog box, SAS EM removed all variables that cannot reach this level. Since we could not apply the second removing rule, the correlation to the target variable, in that dialog box we simply leave the check boxes of "R-squire method" and "Chi-square" unmarked, then SAS EM wound not apply this.

In its variable transformation function, SAS EM first presents the attributes and general descriptive information to the users. The users only need to specify what transformation they need, then SAS EM realize it automatically. For example, when we intend to "maximize the normality" of a specific variable, SAS EM can analyze the its data and choose the best formula to achieve that. Then it tells the users what formula it uses and how the normality is after transformation. The original variable is then set to be not used and a new variable, which is transformed from the original one is put into the data set for data mining.

6.4 Mining Process in SAS EM

In SAS EM, many data mining models can mine one training data. For each model, we need to define which method it uses first, like TRI, LRA or ANN. Then we need to specify the certain criteria it applies. All the specification is determined through dialog boxes. We construct 16 models totally, including 3 LRA (stepwise, forward and backward), 3 TRI (Chi-square, entropy reduction and Gini reduction), 5 ANN models in group one (varied layer numbers and constant hidden unit number) and 5 ANN models in group two (varied hidden unit number and constant layer number).

The results from LRA models are lists of predictors that the model found to be valid. The effective direction (positive or negative) and the effective t-score were also presented.

TRI models generated decision trees. All the trees were clearly shown in diagrams and the corresponding figures (of both the training and validation subset) were specified, so the users can see the proportion of each leave clearly.

Each ANN model gave nothing but a huge weight list of each individual value of the variables. Users cannot see patterns or predictors intuitively from the list table.

6.5 Making Predictions in SAS EM

SAS EM enables a specific step in data mining, named Score, which makes prediction on unknown data after implementing certain mining process on the training data.

Before making a prediction, the user need to indicate which model's result to be used for SAS EM to predict, and on which predicted data set. After predicting SAS EM will add two variables into the predicted data, indicating the possibility of reaching and not reaching an agreement for that corresponding case. (Surely the sum of possibility of reaching an agreement and the possibility of not reaching an agreement is always 1, or 100%). If the predicted possibility of reaching an agreement is greater than 0.5, we regard it as predicting an agreement will be reached; otherwise we take as predicting no agreement. Then we compare them with the actual values of whether an agreement has been reached.

Since SAS EM does not offer proper function to count that, we have to export the prediction table out of SAS system, saved in an Excel 97 format, and then input it into SPSS. Using SPSS we can conveniently count the correct predictions.

7 Findings

Before reporting the findings on predictors, we briefly discuss prediction accuracy. The original data has 53% cases where agreements were reached and another 47% where no agreements were reached. Therefore without further information the prediction accuracy of achieving a compromise cannot exceed 53%.

After applying the data mining results to make predictions, we found that data mining can increase prediction accuracy since their predictions were all above 53%. Because only LRA and TRI methods can identify what are effective predictors, when we try to find the predictors of the negotiation results, we need to see the two methods' prediction accuracy first. In Table 3 the average prediction accuracy of each model are presented.

Table 3. Average prediction accuracy of 5 runs

Logistic regression		Tree and rule induction		Neural network (1-5 layers, 3 units)		Neural network (1-5 units, 2 layers)	
LR_Step	62.40	Tree_Chi	75.14	N_1layer	57.83	N_1unit	57.87
LR_Back	59.28	Tree_Entr	75.33	N_2layer	57.67	N_2unit	59.28
LR_Frwd	61.93	Tree_Gini	75.26	N_3layer	57.76	N_3unit	57.69
				N_4layer	57.53	N_4unit	57.84
				N_5layer	57.63	N_5unit	57.94

* *All figures indicate the percentage of correct predictions.*

From the table we can see that the lowest prediction accuracy is 57.53% and the highest was 75.33%. This indicates that at least some variables may be effective predictors or the groups of variables and their values are good predictors.

7.1 Results from LRA Methods

In Section 5.2.6 we noted that to determine negotiator's behaviors, we would use three LRA methods (stepwise, forward and backward) on five pseudo-randomly generated samples from the database. Prediction accuracy for each of the 15 models is given in Appendix I. From the values in Table 3 we see that the average 15 prediction accuracy of the 3 LRA methods are 62.40, 59.28 and 61.93% respectively. This means that it is possible to increase the prediction accuracy over 5%.

In order to conduct logistic regression SAS EM automatically divided all nominal variables into many binary variables. A binary variable was defined for each value of the nominal variable. For example, the variable OFF_1 (number of offers sent on the day before deadline), which took values 0, 1, 2 3 and 4, was divided into four binary variables: OFF_1 1, OFF_1 2, OFF_1 3, OFF_1 4. The last figure of each variable indicated the number of sent offers. Therefore we are able to see more details that each value's impact is also identified. However, since the variable OFR has too many values and we mainly concern the variable's general impact rather than that of each value, it is treated as a continuous variable (called interval variable inside SAS EM). It regards the variable Gender as a nominal variable since its missing value is indicated as N/A but not just a blank. The male participant's always-stronger negative impact on reaching an agreement reveals that males are somewhat unlikely to compromise than females.

The results of LRA models are list of identified effective predictors. The effective T-score of each predictor is also indicated. The identified predictors are ranked according to their T-score's absolute values. (The predictors obtained from LRA models are given in Table 4.) Because the training and validation data varied each time, the absolute values of the T-scores of individual models were different for each run.

In this thesis we concern the behaviour patterns rather than the accurate prediction impact value of each predictor. Therefore, we only need to find out the effective predictors and then we can see the patterns. We do not need to analyze the detailed values of their T-scores. In the five rounds of LRA data mining analysis, if a variable was identified to be effective predictors several times and its T-score is not tiny, then we can believe that it is effective. While knowing its impact direction (positive or negative to the target) we are able to generate the basic patterns.

Each of the three different LRA model as well as models of the same type but constructed for a different subset of data yields different results. This is because the training data was different in each case or different methods were used in each LRA type. For each model from the fifteen that were constructed, however, the prediction accuracy exceeds the minimum of 53% because the predictors are valid. The list of all predictors identified is given in table 4. For each predictor on the list (1) the method(s) used to determine, (2) the rank position of its effective T-score value (the predictor with largest T-score ranks the first, and that with the smallest T-score ranks the last. So the bigger the value of ranking position is, the less impact it has) and (3) the effect direction (negative or positive to the target) are given.

Table 4 Predictors Found by LRA Models

Round	Model	OFF_2 1	OFF_2 2	OFF_4 1	Gender M	Gender F
1	LR-step					
1	LR-forwd					
1	LR-back			29		
2	LR-step				2-	3-
2	LR-forwd	20-	22-	17-	26-	30-
2	LR-back	3-	21			
3	LR-step	7-	10-			
3	LR-forwd	7-	10-			
3	LR-back	3-	11-	4-	9-	12-
4	LR-step					
4	LR-forwd					
4	LR-back					
5	LR-step	5-	4-		2-	6-
5	LR-forwd				23-	30-
5	LR-back					
TOTAL		6	6	3	5	5

Round	Model	OFR	OFF_ALL	OFF_1 1	OFF_1 2	OFF_1 3	OFF_1 4
1	LR-step	5	4	10-	12-	9-	
1	LR-forwd	5	4	10-	12-	9-	
1	LR-back						

Round	Model	OFR	OFF_ALL	OFF_1 1	OFF_1 2	OFF_1 3	OFF_1 4
2	LR-step		1	5-	4-	6-	
2	LR-forwd	18		29-	24-	28	
2	LR-back		24				
3	LR-step	3		11-	9-	6-	5-
3	LR-forwd	3		11-		6-	5-
3	LR-back	7		1-	2-		
4	LR-step		1	2-	3-	6-	
4	LR-forwd		18	17-	21-	23-	
4	LR-back			1-	20-		
5	LR-step	1		8-	12-		7-
5	LR-forwd	24	28	14-		25-	27-
5	LR-back						
TOTAL		8	7	12	10	10	4

** Figures inside the table indicate the rank of the effective T-score. The higher it ranks (the less the figure is), the less impact it has. If it has a negative impact, then a "-" symbol is marked behind the figure.*

7.2 Patterns Obtained with LRA Methods

7.2.1 "Last-minute" offers

Several findings are listed below:

- **If the negotiations are ongoing one day before the deadline then it is less likely to reach an agreement than the other parties. The probability of reaching an**

agreement is significantly reduced if still negotiating on the last day before the deadline.

Negotiators' behaviour during the last day before deadline can be used to predict the result. The binary variables OFF_1 1, OFF_1 2, OFF_1 3 and OFF_1 4 that were created from the variable OFF_1 were identified 12, 10, 10 and 4 times representatively to be effective predictors.

They received the highest observation frequencies as effective predictors among all other variables. Since the parties still sent offers on the last day, they continued their efforts in reaching an agreement. Sending an offer on the last day may be considered as exerting pressure, made by at least one side to complete the negotiation successfully. However we observed that the probability of success of this pressure is low. That is, "last-minute" offers tend to be rejected. Therefore it is reasonable that the sending of offers on the last day has some relation with out target variable. Furthermore we found out of the offers sent on the last day, all, except for one very small positive T-score, had a negative effect to the target. That meant that there was less possibility of reaching an agreement than the average if the parties kept negotiating till the last day.

When we looked at the population - the whole 3050 cases – this pattern was verified. Among the 3050 cases, 268 cases (9.79%) sent one or more offers on the last day. Among them 80 reached agreements. That is 29.85% only, almost half of the average possibility. Therefore the first result from LRA models was verified. This finding tells us that the participants of Internet negotiation should try to avoid waiting to the last day if they wish to reach agreement, since the possibility of reaching agreement is reduced by almost 50%.

7.2.2 The impact of communication

- **The more offers sent, the greater the probability that the agreement will be reached.**

Among the 15 LRA data mining results, 8 times the variable OFR (total number of offers sent) was found to have positive effect on the target. In six models OFR was identified as a strong predictor (ranked 1st, 3rd, 3rd 5th, 5th, and 7th, see table 4) and in two models this variable was found a weak predictor, and all their impact directions are positive. Therefore we could conclude that the more offers sent, the greater the probability that the agreement will be reached during Internet negotiation. This is a communication effect, since that the exchange of offers means that the party is interested in achieving an agreement.

- **Number of offers sent 4 days before the deadline has positive effect on the target.**

Another predictor we found was the variable OFF_ALL (number of offers sent 4 days before deadline). Seven times, the LRA methods found that it had strong positive effect on reaching agreement, and four times as a strong predictor (ranked in 1st, 1st, 4th and 4th separately). The variables of the Inspire system were defined to pay special attention to the behavior at the final stage. If we call the last 4 days the final stage of this negotiation (and the days before the last four days the early stage and middle stage), then there were 3 variables used to record the offers sent for the last stage including OFF_1. The other two were OFF_2 (number of offers sent on 2 days before deadline) and OFF_4 (number of offers sent on 4 days before deadline). Therefore OFF_ALL recorded the number of offers sent at the early and middle stage.

The strong positive effect of OFF_ALL indicated that the more offers sent before the final stages, then the more probability that the agreement would be reached. We also found that OFF_2 and OFF_4 generally had negative effect (see table 4). This confirmed that in its early and middle stages, the number of offers sent had positive impact on reaching agreements, because we have known that the total number of offers sent (OFR) had a positive relation with the target, and the offers sent in the final stage (OFF_1, OFF_2 and OFF_4) had a negative effect. Therefore, the offers sent in the early and middle stages (OFF_ALL) definitely had a positive impact on the target. This means that early engagement in negotiation is important. It might also be that in the early stages of the negotiation there is not much pressure to achieve a compromise and the parties may concentrate on learning about each other and understating their needs and preferences. It is good for negotiators to avoid waiting too long.

7.2.3 Gender and agreements

- **Males are somewhat less likely to reach agreements than females.**

There were totally 5 times those LRA methods identified genders as an effective predictor. We found that Gender_M and Gender_F were always identified together and no one was observed effective without another's appearance. Further analysis revealed that Gender_M always had stronger negative effect on target than Gender_F, but the difference was not big. (Please refer to table 4). Therefore we may say that male participants are somewhat more difficult in reaching agreement than females.

7.3 Patterns Found by TRI Methods

7.3.1 Reduction of decision tree models

We have found that all the data mining methods can increase prediction accuracy. The TRI methods' prediction accuracies ranged from 72.563% to 77.805% on the negotiation results. In the source data, 53.049% of the cases reached an agreement and 46.951% did not. Without further information then, the prediction accuracy of whether the agreement can be reached cannot be higher than 53.049%. This indicates that the patterns obtained with TRI methods can be effectively used to predict the target variables and increase prediction accuracy by over 19% on average. .

Using TRI methods we have obtained 15 different decision tree models. All these models could be grouped into two types based on the "trunks". Because there are two different "trunks" that all the models contain one of them. When we compared the prediction accuracy of the trees having the "trunk only" and the "trunk and branches" trees, we obtained no significant difference between them because the p-value that they were the same is 0.744 (by Games-Howell Post-Hoc ANOVA analysis method). This implies that using truncated models (e.g., only trunks of the trees) predictions are as effective as using complete models. While such "branch variables" as CBORN, CGUESS, FIRSTLAN…etc, have some impact on prediction accuracy, this impact is very small.

7.3.2 Decision Tree No.1

Decision Tree No.1 develops its first level of leaves according to the variable OFR (total number of offers sent). Please refer to Figure 1.

7.3.2.1 Decision Rule 1-1:

- **IF during the negotiation the number of offers sent by one party is smaller than two**
- **THEN there is very low (i.e., 0.119) probability that the parties achieve an agreement.**

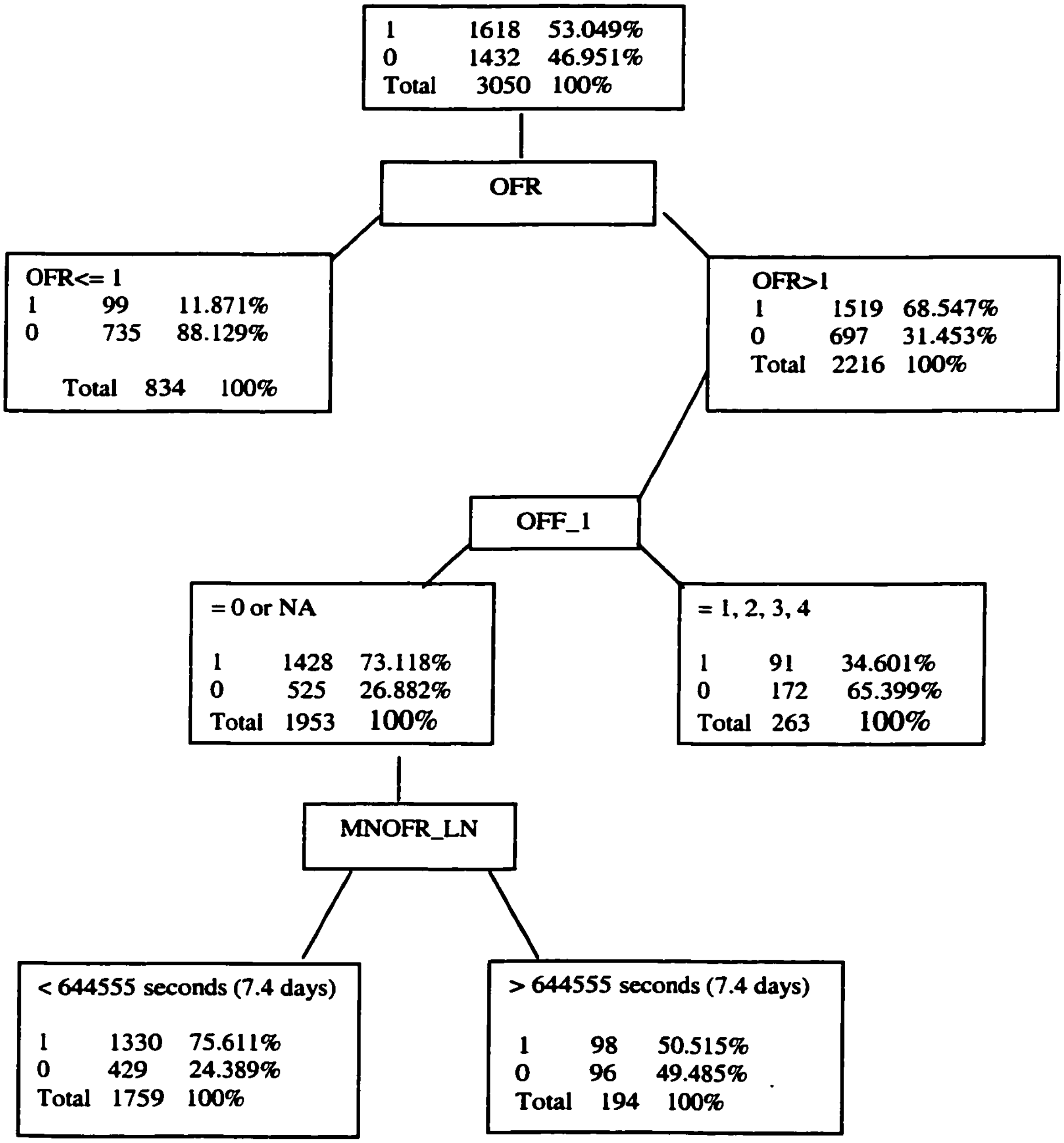
Using the SAS EM we obtained that almost one third of the population (i.e., 27.34%, 834 cases) sent less than two offers. From this group only 11.871% (99 cases) reached agreements. If the negotiator sent two or more offers, then we can use predictor OFF_1 representing the number of offers sent 1 day before negotiation deadline.

7.3.2.2 Decision Rule 1-2:

- **IF the negotiator sent more than one offer, and**
- **IF this negotiator sent also at least one offer on the last day before the deadline**
- **THEN there is low (i.e., 0.346) probability that the parties achieve an agreement.**

We found that 263 cases (8.62%) among the population have offer(s) sent one day before the deadline. Among them only 91 cases (34.601%) have reached agreements, which is much lower than the average (53.049%). In most cases (91.279%), the negotiators did not send offers on the last day and their possibility of reaching agreements

Figure 1. Decision Tree No.1



was up to 73.118%. Therefore we could make the conclusion that sending offers on the last day had a low possibility of reaching agreements (34.601%). The TRI models reveal the same conclusion as that from LRA models.

The results obtained from TRI models also show indicate that if more than 1 offer was sent and no offers sent on the last day, the time between the offers could affect the result. Then we have two more decision rules:

7.3.2.3 Decision Rule 1-3:

- **IF more than 1 offers are sent, and there is no offer sent on the last day before deadline, and**
- **IF the mean interval time between offers is less than 7.44 days**
- **THEN there is a high probability of reaching an agreement (75.611%).**

The results show that there are 1759 cases (57.672%) that more than 1 offers are sent, no offer is sent on the day before deadline and the mean interval time is less than 7.44 days. Among them there are 1330 cases (75.611%) that the negotiators reach agreements.

7.3.2.4 Decision Rule 1-4:

- **IF more than 1 offers are sent, no offer is sent on the last day before deadline, and**
- **IF the mean interval time between offers is longer than 7.44 days**

- **THEN there is only an approximately average probability of reaching an agreement (50.515%).**

The results show that there are 194 cases (6.36%) that more than 1 offers are sent, no offer is sent on the day before deadline and the mean interval time is longer than 7.44 days. Among them there are 98 cases (50.515%) that the negotiators reach agreements.

Therefore we could say that if the offers sent have a long interval time between offers, the possibility of reaching agreements could be significantly reduced. The negotiators, after sending one offers, should avoid waiting too long (more than a week) to send next one, otherwise the possibility of reaching agreements is reduced.

In general Tree No.1 indicates 3 behavior patterns in Internet negotiation:

- **Few offers sent, very low reaching agreement possibility;**
- **Even more than one offers sent, waiting to last days can lead low possibility;**
- **The interval of time between offers should not be too long; if exceeding one week, the probability was reduced a lot.**

7.3.3 Decision Tree No.2

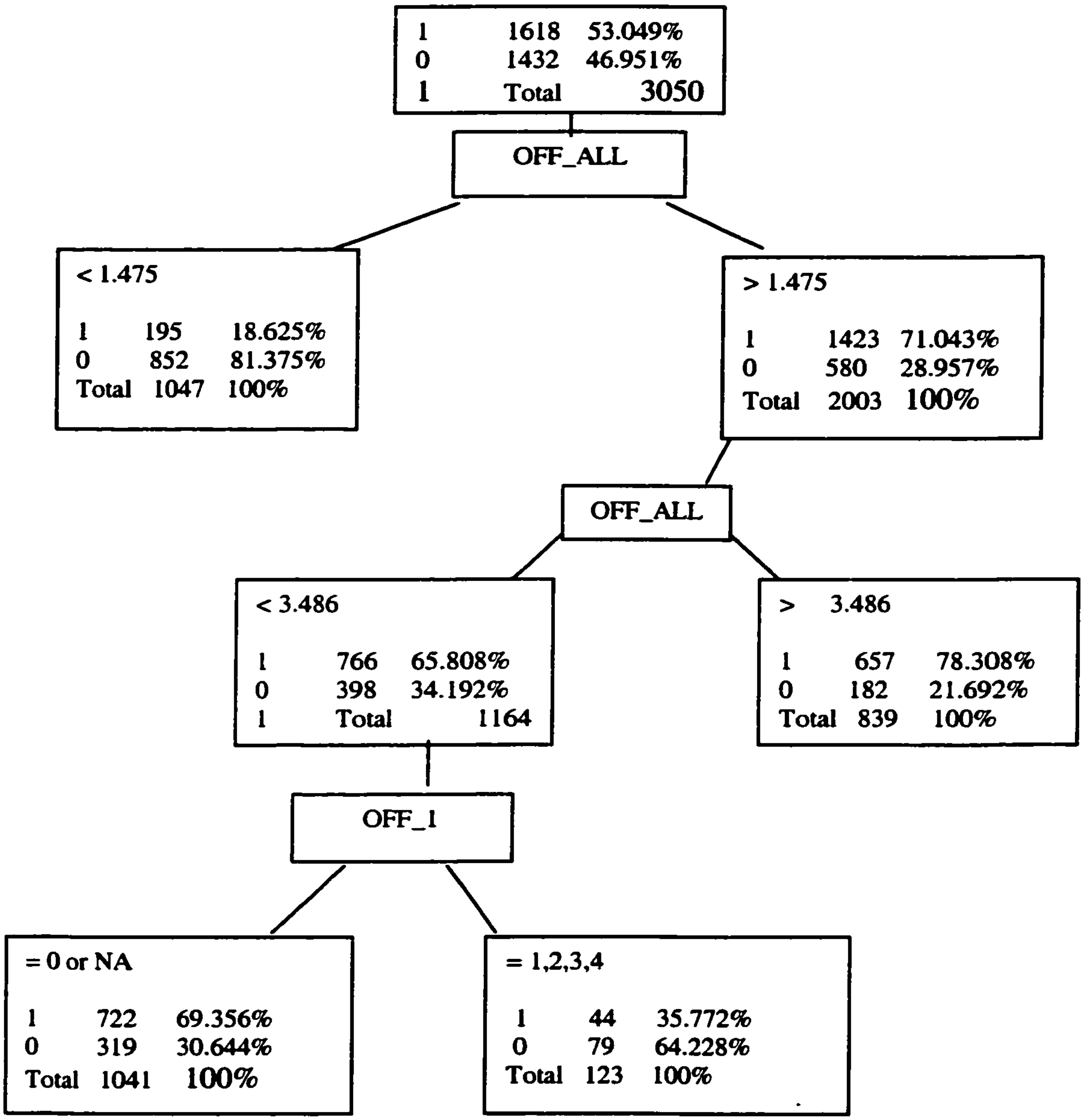
The second decision tree first concerned the variable OFF_ALL (the number of offers sent before the final stage, that was, at least 5 days before the deadline). Please refer to figure 2.

7.3.3.1 Decision Rule 2-1:

- **IF none or only one offer is sent four days before the deadline**

- **THEN the possibility of reaching an agreement is low**

Figure 2. Decision Tree No.2



Among the population there are 1047 (34.328%) cases that have none or only one offers sent 4 days before the deadline. Only 195 (18.625%) of them have reached agreements. The probability is quite low.

7.3.3.2 Decision Rule 2-2:

- **IF more than 3 offers are sent four days before the deadline**
- **THEN the probability of reaching agreements is quite high (78.380%).**

There are 839 cases (27.508%) among the population that have 4 or more offers sent 4 days before the deadline, while they have 657 cases (78.380%) with agreements reached.

This indicates that to send more offers 4 days before the deadline can lead high probability of reaching an agreement.

There are 1164 cases (38.066%) of the population that have 2 or 3 offers sent four days before the deadline; among them 766 cases have reached agreements. The possibility is 65.808%. One more predictor, OFF_1 (number of offers sent on one day before the deadline,) is identified to be effective.

7.3.3.3 Decision Rule 2-3:

- **IF 2 or 3 offers are sent four days before the deadline, and**
- **IF no offer is sent on the day before the deadline**
- **THEN there is approximately 70% probability of reaching an agreement.**

7.3.3.4 Decision Rule 2-4:

- **IF 2 or 3 offers are sent four days before the deadline, and**
- **IF there is offer(s) sent on the day before the deadline**
- **THEN there is only 36% probability of reaching an agreement.**

If 2 or 3 offers are sent in the early and middle stages, but get no agreements before the deadline, then sending offers on the last day has only a 35.772% chance to reach an agreement. That is much less than the average.

Tree No. 2 indicates the following behavior patterns of Internet negotiations:

- **Number of offers sent four days before the deadline has positive impact on the probability of reaching agreements;**
- **Sending offers in the final stage has low possibility of reaching agreements.**

7.3.4 Comparison of Patterns found by LRA and TRI Models

By comparing the results from LRA and TRI models, we found most of the patterns they found are the same. If we define the last four days before the deadline as the final stage of negotiation, then before that we can call it as the early and middle stage of negotiation. The patterns are generated in Table 5.

Exchange of offers is an obvious requirement. What is less obvious is that the success of the Internet negotiation also depends on the number and timing of these exchanges. One possible explanation is that the narrow bandwidth of the Internet as opposed to the face-

Table 5. Behaviour Patterns of Internet Negotiations

No.	Pattern Description	Found by LRA Models	Found by TRI Models
1	If the parties continue negotiations on the last day before deadline, the probability of reaching a compromise decreases.	Yes	Yes
2	The more offers the parties exchange the greater the probabiiity that they will reach an agreement.	Yes	Yes
3	The offers sent in early stage of negotiation have stronger impact on reaching an agreement than the offers sent at the later stage.	Yes	Yes
4	Males are somewhat less likely to reach agreements than females.	Yes	No
5	If the time between offers, that the parties exchange, exceeds one week, the possibility of reaching agreements is significantly reduced.	No	Yes

to-face negotiation requires compensation in the volume of communication. This may also conform to the need for early exchanges and the negative impact of the exchanges just before deadline. The parties may need to get acquainted and may be able to change their perspectives without being pressured to reach a compromise at the last moment.

This has probably more to do with the relationship aspects of the negotiation than the negotiation problem. There are neither cognitive nor technical difficulties in the understanding and control of the problem process. The problem is fairly simple and well structured, and the analytic and process visualization tools allow evaluating and controlling both the individual issues and the overall process.

7.4 Comparison of Prediction Accuracy

The prediction accuracy is the main concern for users in real life. After implementing the 3 data mining methods we obtained 5 prediction accuracy for each of the 16 models of the 3 methods. Since each time all the models were mining the same training and validation data and predicting the same score data, then our comparison of their accuracy is valid.

We adopted the Games-Howell Post Hoc Test ANOVA (single factor with blocks) to compare their means. ANOVA is the most popular statistical method to compare means of different variables. The Games-Howell test ANOVA analysis does not require the assumption of equal variance of each variable. Therefore we adopted it and we did not need to test their variances. This method was offered by the SPSS 10.0 version. (Please refer to Appendix II and III to see the One-Way ANOVA results. The p-value indicates the possibility of the statement that the means were the same.)

7.4.1 Comparison of LRA Models Prediction Accuracy

We found that we could almost surely conclude that:

- **The LRA stepwise and LRA forward models had the same prediction**

accuracies (P-value is 1.0).

At 85% confidence level we could conclude that:

- **LRA stepwise and LRA forward models had better prediction accuracies than that of the LRA backward model (P-value is 0.081 and 0.144 separately).**

LRA stepwise (its prediction accuracy is average 62.39%) and LRA forward (average 61.93%) have 3.11% and 2.65% more correct predictions separately on average, while LRA backward can make an average 59.28% correct predictions.

7.4.2 Comparison of TRI Models Prediction Accuracies

We are confident that:

- **All three TRI models have same prediction accuracies**

The p-values that they are the same are all 1.0. No model is more accurate at prediction than others. The TRI Chi-square, TRI entropy reduction and TRI Gini reduction have an average 75.14%, 75.33% and 75.14% prediction accuracy separately.

7.4.3 Layer Number's Effect on Prediction Accuracy in ANN Models

From the 25 figures (5 for each model) we cannot see any prediction accuracy difference between any two models from model 7 to model 11 (p-values were all equal to 1). Their average prediction accuracy is 58.03%, 57.67%, 57.76%, 57.52% and 57.63% representatively. We may conclude that:

- **The layer number has no impact on the prediction accuracy, though each**

model differs slightly from the others.

7.4.4 Hidden Unit Number's Effect on Prediction Accuracy in ANN Models

From the 25 figures (5 for each model) we cannot observe a difference in prediction accuracy between any two models from model 12 to model 16 (p-values all equaled to 1.0). Their average prediction accuracy is 57.87%, 57.69%, 57.82%, 57.84% and 57.94% representatively. We might conclude that:

- **The hidden unit number has no impact on the prediction accuracy though each model made a slightly different prediction each time.**

7.4.5 Prediction Accuracy Difference between LRA, TRI and ANN Methods

In order to have a general idea of how accurate each data mining methods can make prediction, we grouped all the 16 models into 3 groups according to the data mining methods it used. Then we adopted One-Way ANOVA to compare their means (See Appendix III). We observe these three methods provide different prediction accuracies since all the p-values of the hypothesis that they were the same were 0, that is their prediction accuracy is not the same. There were significant differences between each other. TRI methods have the highest mean prediction accuracy: 75.20%. LRA methods rank the second with a mean 61.2% and ANN methods have lowest mean prediction accuracy of 57.78%.

By considering the paired comparisons of 16 models, we found all 3 TRI methods having significantly higher accuracy than all the other 13 models, while there was no difference between them. LRA stepwise and forward models had higher prediction accuracies than the remaining 11 models and they were the same at an 85% confidence level. There is no significant difference between the LRA backward model and other 10 ANN models, but all the 10 ANN models have same prediction accuracy. The sequence of PA (prediction accuracy) is as follows:

$$\mathbf{PA}_{\text{TRI models}} > \mathbf{PA}_{\text{LRA stepwise}} = \mathbf{PA}_{\text{LRA forward}} > \mathbf{PA}_{\text{LRA backward}} >= \mathbf{PA}_{\text{ANN models}}$$

7.5 Computation time of LRA and TRI Models

All computation was done on the Hewlett-Packard HP71 personal computer with Intel Celeron 533 MHz CPU and 128 MB RAM. Although we planed to measure the mining time of all models, we were not able to achieve this with respect to the ANN models. This is because the SAS EM often stopped running the ANN model while it had more than 1 layer during the mining process. The stop could not be noticed until we found the process was abnormally long and then we manually resumed it. This might be caused by this specific software design defect. Therefore we failed to collect reliable time consumption data for ANN models with more than 1 layer. Only the ANN model had 1 layer and 3 hidden units did not stop. Therefore we could only compare this ANN model with other LRA and TRI models.

The One-Way ANOVA Post-Hoc Test of the 7 models (last 9 models were not included) revealed that (please refer to Appendix IV):

- The time consumption between LRA stepwise, LRA backward and LRA forward models are the same. (P-values were1.0, 0.997 and 1.0 separately.)
- The time consumption between TRI Chi-square, TRI entropy reduction and TRI

Gini reduction models are the same. (P-values were0.997, 0.988 and 1.0 separately.)

- The ANN 1 layer 3 hidden unit model consumed the same time as 3 LRA models. (P-vales were 1.0,1.0 and 0.987.) It is longer than TRI models at 95% confidence level. It takes more than one and a half minutes for this ANN model than each TRI model.
- When we compared the time consumption of the 3 data mining methods after grouping these 7 models by their data mining methods, the ANOVA analysis showed that: (1) TRI methods consumed the least time with the p-values 0.0 and 0.0; (2) At 85% confidence level we believed that LRA models and ANN one layer three hidden units consumed the same processing time (P-value 0.872).

The sequence of time consumption (TC) is as following:

$$TC_{\text{TRI models}} < TC_{\text{LRA Models}} = TC_{\text{ANN 1 layer 3 hidden units}}$$

7 Benefits and Limitations

8.1 Benefits

This study's main research target was to find out some behaviour patterns that were difficult to identify by traditional statistical methods. A better understanding of the factors that affect the negotiation result can increase the Internet negotiators' work efficiency. Any failed negotiation is usually just a waste of time and energy for both parties and a waste of Internet resources. Therefore our findings from this study can benefit Internet negotiators directly by improve their negotiation efficiencies because the patterns we found using the data mining methods can be easily understood and applied during Internet negotiations. There are 3 valuable guidance rules, which we believe should receive sufficient attention so that the chance of reaching agreement will be largely increased:

- **Avoid waiting to send offers till the last several days. There is little chance to reach agreement at the last minute.**
- **Send more offers if possible, especially before the final stage. It can significantly increase the probability of reaching agreements.**
- **Avoid a long time interval between offers. The interval of more than a week would greatly decrease the chance of reaching agreements.**

These 3 rules are easy for negotiators to apply. Because there are more and more people negotiating through the Internet, increasing the negotiation efficiency can bring more benefit to the negotiators and reduce their cost by saving their time, human resources and other communication expenses.

This study can also benefit researchers who are interested in Internet negotiation or other Internet communication studies. Our findings can be used to compare with other Internet communication behaviors to help them construct new hypothesis/models on Internet negotiations.

The study results of different data mining methods' prediction accuracy can help the users to select the most suitable method. Suitable data mining methods can give users more accurate predictions on their interested target. A better prediction of the target certainly can increase their benefits or reduce their costs. Our results showed that the TRI methods could make the highest prediction accuracies while consuming the least processing time; therefore it is a good choice for users.

The study of data mining's prediction accuracy can also benefit data mining developers. After knowing the differences and continuing research on the reasons for the differences, they can therefore improve current data mining methods/software and generate guidance to assist users to choose the most suitable methods.

8.2 Limitations

7.4.6 Generalization

Since this study is a case study, concerns about generalization remain legitimate. First, data mining regards the data base as the population, but the data from Inspire system may not be the same as other Internet negotiations. Therefore the data mining results on this data (though mined on the whole population here) may not necessarily be applicable to all cases.

Secondly, our data mining prediction accuracy analysis is limited in this data set. For other data set - especially for the set containing different variable type like nominal or interval targets or other data size - the results may be different. Therefore more data mining prediction accuracy studies should be conducted to get more generalizable results.

7.4.7 Validity

To enhance both internal and external validity, the following consideration were incorporated into this research:

- Freedom from sampling bias was avoided since we applied data mining to the whole population. (This population may have a bias, but it is out of our control.)
- All the data mining was applied on the same training and validation data set and scored on same data set; therefore the comparisons of prediction accuracies are valid.
- Collecting results from 5 repeated processes, and analyzing the results with the ANOVA method, gave us valid comparison results.
- This specific data mining software might limit the results; therefore its external validity is reduced.
- The 3050 data collected represents 1525 negotiation cases. The validity may be reduced since two parties of the same case are analyzed together. This is because that the data were collected without separating them into two tables, and we do not know who negotiated with whom.
- The results obtained are valid to the current data in the Inspire system.

7.4.8 Reliability

To ensure reliability, each time we implemented all 16 models on the same training and validation data set to compare the results and scored the same data each time to compare their prediction accuracies; then the same process was repeated by 5 times. Similar results (both on the Internet behaviour patterns and the prediction accuracies) were obtained each time. It proved that the reliability of the study is good. By applying the rules mined out from training and validation data set, back to the population and we observed that the rules were still effective, and knew that the rules were reliable and valid.

8.3 Potential Future Research

Since there were some limitations to this study, more research should be conducted to check the external validity of this enquiry. For example, by collecting data on Internet negotiations through other methods rather than the Inspire system, researchers can then construct and verify the hypothesis and models based on the rules found in this study. Applying more data mining methods on different types and sizes of data can complement our study in order to obtain more information on the applications of different data mining methods.

9 References:

Adler, N. J., R. Brahm, et al. (1992). "Strategy Implementation: A comparison of face-to-face negotiations in the People's Republic of China and the United States." Strategic Management Journal **13**(6): 449-466.

Adriaans, P. and D. Zantinge (1996). Data Mining. New York, Addison-Wesley.

Breiman, L., J.H. Friedman, et al. (1984). Classification and Regression Trees. Pacific Grove: Wadsworth.

Bui, T. (1994). "Designing multiple criteria negotiation support systems: Frameworks, issues and implementation". Tenth Internal Conference: Expand and Enrich the Domains of Thinking and Applications, Springer-Verlag.

Carroll, M. L. (1996). Cyberstrategies: How to build an Internet-based information system.Van Nostrand Reinhold Publishing.

Chatterjee, K. and G. L. Lilien (1984). "Efficiency of alternative bargaining procedures: An experimental study." Journal of Conflict Resolution **28**(2): 270-295.

Chen, M.-S., J. Han, et al. (1996). "Data Mining: An overview from a database perspective." IEEE Transactions on Knowledge and Data Engineering. **8**(6): 866-883.

Drake, L.E. (1995). "Negotiation Styles in Intercultural Communication". International Journal of Conflict Management. **6**(1): 72-90.

Fayyad, U. M. (1996). "Data Mining and Knowledge Discovery: Making Sense Out of

Data" IEEE Expert. **11**(5): 20-25.

Fayyad, U.; G. Piatetsky-Shapiro; and P. Smyth (1996). "From Data mining to knowledge discovery in databases." AI Magazine **23**(8):37—54.

Glymour, C., D. Madigan, et al. (1997). "Statistical Themes and Lessons for Data Mining." Data Mining and Knowledge Discovery. **1**(1): 11-28.

Graham, J. L., A. T. Mintu, et al. (1994). "Explorations of Negotiation Behaviours in Ten Foreign Cultures Using a Model Developed in the United States." Management Science. 40(1): 72-95.

Guttman, R. H., A. G. Moukas, et al. (1998). "Agent-mediated electronic Commerce: A Survey." Knowledge Engineering Review. **13**(3).

Harris-Jones, C. (1997). Data Mining Tools Survey. London, AMS Center for Advanced Technologies.

Hendon, D. W., R. A. Hendon, et al. (1996). Cross-cultural negotiations. Quorum Books.

Henery, R.J., D. Michie, D.J. Spiegelhalter, and C.C. Taylor (eds.). (1994). Machine Learning, Neural and Statistical Classification. New York: Ellis Horwood, pp. 6—16.

IBM (2000). Data Mining. IBM corporation, http://www.ibm.cam/datamining

IIASA (1998). http://www.iiasa.ac.at/Research/TAP

James, K. (1985). *Classification Algorithms*. Thompson Publishing, pp123-125.

Jain, B.A., and J.S. Solomon (2000). "The Effect of Task Complexity and Conflict Handling Styles on Computer-supported Negotiations." Information &. Management. Amsterdam. **37**(4):161-168.

John, G. H. (1997). Enhancements to the Data Mining Process. Computer Science Department, School of Engineering, Stanford University: pp194.

Kersten, G.E., (1985), "NEGO - Group Decision Support System", Information and Management, **8**(5), p. 237-246

Kersten, G. E., and S. J. Noronha (1999). "Negotiation Via the World Wide Web: A Cross-Cultural Study of Decision Making" Group Decision and Negotiation. 8: 251-279.

Kersten, G. E., and S. J. Noronha (2000). Modeling Integrative and Distributive Negotiations. Interneg, INR03/00, http://interneg.org/interneg/research/papers/. Accessed: June 2000.

Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection". International Joint Conference on Artificial Intelligence (IJCAI).

Lax, D. and J. K. Sebenius (1986). The manager as negotiator: Bargaining for cooperation and competitive gain. New York, The Free Press.

Liang, T.P. (1992). "A composite approach to inducing knowledge for expert system design." Management Science. 38:1—17.

Limb, P. R. and G. J. Meggs (1994). "Data Mining - tools and techniques." *BT*

Technology. **12**(4): 32-41.

Lo, G and G. E. Kersten (1999). "Negotiation in Electronic Commerce: Integrating Negotiation Support and Software Agent Technologies." Proceedings of the 5th Annual Atlantic Canadian Operational Research Society Conference.

Mahadevan, A. (1999). "Data Mining in Social Research: A Case of Modeling Anonymous Negotiation". Thesis to the Faculty of Graduate Studies and Research, Carleton University, Ontario.

Mahadevan, A. and Ponnudurai K. (1999). "Knowledge Discovery in Database and Decision Support." http://interneg.org/interneg/research/papers/1999/04.pdf

Messier, W.F., and J.V.Hansen (1988). "Inducing rules for expert system development: and examples using default and bankruptcy rules." Management Science. pp.1403 – 1415.

Nyhart, J.D. and C. Goeltner (1997). Computer Models as Support for Complex Negotiations. International Conference for the Society for General System Research, Budapest.

Pregibon, D. (1998). "Data Mining: What's at stake for the field of statistics". Talks in Statistics. Statistics Canada, Ottawa, Canada.

Quinlan, J.R. (1993). C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann Publishers.

Rangaswamy, A. and G. R. Shell (1994). "Using Computers to Realize Joint Gains in Negotiations: Toward an Electronic Bargaining Table". Computer-Assisted Negotiations and Mediation Symposium. Harvard Law School, Cambridge, MA.

Romesburg, H. C. (1990). Cluster Analysis for Researchers. Malabar, Robert E. Krieger Publishing Company.

Sandholm, T. (1999). "Automated Negotiation. The Best for All Concerned." Communication of the ACM. **42** (3):84-85.

SAS (1999). Data mining. SAS Institute Inc. 1999.

SAS (2001) Help File. Version 8.1

SAS (2001) On-line Documentation Version 8.1

Sebenius, J.K. (1984). Negotiating the Law of Sea. Cambridge. MA. Harvard University Press.

Shavlik, J.W., and T.G. Dietterich (1990). Readings in Machine Learning. San Mateo, CA: Morgan Kaufmann, pp. 1—10.

SPSS (1998). SPSS in Data Mining. Chicago, SPSS Inc. http://www.spss.com.

Spangler, W.E.; J. H May; L.G. Vargars (1999). "Choosing data-mining methods for multiple classification: Representational and performance measurement implications for decision support." Journal of Management Information Systems. Armonk. **16**(1):37-62.

Strobel. M. (1999). On Auctions as the Negotiation Paradigm of Electronic

Markets. Zurich, BM Zurich Research Laboratory: 7.

Tessmer, A.C.; M.J.; Shaw and J.A. Gentry (1993). "Inductive learning for international financial analysis: a layered approach." Journal of Management Information Systems **10**(2):17—36.

Tukey, J. W. (1977). Exploratory Data Analysis. Reading, Addison-Wesley.

Weiss, S. M. and N. Indurkhya (1998). Predictive data mining: A practical guide. San Francisco, Morgan Kaufman Publishers.

Weiss, S.M., and C.A. Kulikowski (1991.) Computer Systems That Learn. San Francisco: Morgan Kaufmann Publishers.

Wirth, R., C. Shearer, et al. (1997). "Towards process-oriented tool support for Knowledge Discovery in databases". Principles of Data Mining and Knowledge Discovery: First European Symposium. J. Komorowski and J. Zytkow, Springer-Verlag.

Appendix I Prediction Accuracies of 16 Models

Table 1. Prediction Accuracies of LRA and TRI models

Round	LR-STEP	LR-BACK	LR-FORWD	TREE-CHI	TREE-ENTR	TREE-GINI
1	61.261	57.903	61.261	75.921	77.805	75.921
2	61.016	58.641	60.360	75.840	75.840	75.840
3	63.227	60.115	63.227	74.775	75.758	75.184
4	62.735	60.688	62.490	76.413	74.693	74.693
5	63.718	59.050	62.326	72.727	72.563	74.038
Mean	**62.391**	**59.279**	**61.933**	**75.135**	**75.332**	**75.255**

* All prediction figures refer to the percentage of correct predictions.

Table 2. Prediction Accuracies of ANN Models with Constant Hidden Unit Number

Round	NEU1LAY	NEU2LAY	NEU3LAY	NEU4LAY	NEU5LAY
1	57.740	57.002	56.921	56.593	57.740
2	57.084	57.084	56.429	56.020	56.511
3	58.804	58.395	58.559	58.722	58.640
4	57.411	56.921	57.248	57.166	56.675
5	59.132	58.968	59.623	59.132	58.559
Mean	**57.834**	**57.674**	**57.756**	**57.527**	**57.625**

* All prediction figures refer to the percentage of correct predictions.
*The hidden unit number in all layers is 3.

Table 3. Prediction Accuracies of ANN Models with Constant Layer Number

Round	N1UNIT	N2UNIT	N3UNIT	N4UNIT	N5UNIT
1	58.477	58.722	58.559	58.395	58.722
2	57.412	56.593	57.002	57.248	57.002
3	57.903	57.084	57.576	58.067	57.412
4	57.985	58.640	58.231	58.313	58.640
5	57.576	57.412	57.740	57.166	57.903
Mean	**57.871**	**59.279**	**57.690**	**57.838**	**57.936**

* All prediction figures refer to the percentage of correct predictions.
*The Layer number in all models is 2.

Appendix II ANOVA Results of 16 Models' Prediction Accuracies

(I) FACTOR1	(J) FACTOR1	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	3.11200	.68914	**.081**	-.30391	6.52791
	3	.45860	.68914	**1.000**	-2.96139	3.87859
	4	-12.74380(*)	.68914	**.000**	-16.73739	-8.75021
	5	-12.94040(*)	.68914	**.000**	-17.89887	-7.98193
	6	-12.74380(*)	.68914	**.000**	-15.86684	-9.62076
	7	4.35720(*)	.68914	**.009**	1.17756	7.53684
	8	4.71740(*)	.68914	**.006**	1.49625	7.93855
	9	4.63540(*)	.68914	**.014**	.94035	8.33045
	10	4.86480(*)	.68914	**.011**	1.10019	8.62941
	11	4.76640(*)	.68914	**.006**	1.48478	8.04802
	12	4.52080(*)	.68914	**.011**	1.36348	7.67812
	13	4.70120(*)	.68914	**.006**	1.47378	7.92862
	14	4.56980(*)	.68914	**.008**	1.47680	7.66280
	15	4.55360(*)	.68914	**.008**	1.45888	7.64832
	16	4.45560(*)	.68914	**.007**	1.34804	7.56316
2	1	-3.11200	.68914	**.081**	-6.52791	.30391
	3	-2.65340	.68914	**.144**	-5.95556	.64876
	4	-15.85580(*)	.68914	**.000**	-19.78346	-11.92814
	5	-16.05240(*)	.68914	**.000**	-20.99238	-11.11242
	6	-15.85580(*)	.68914	**.000**	-18.80592	-12.90568
	7	1.24520	.68914	**.804**	-1.77739	4.26779
	8	1.60540	.68914	**.563**	-1.46751	4.67831
	9	1.52340	.68914	**.791**	-2.08260	5.12940
	10	1.75280	.68914	**.668**	-1.92873	5.43433
	11	1.65440	.68914	**.560**	-1.48972	4.79852
	12	1.40880	.68914	**.505**	-1.53433	4.35193
	13	1.58920	.68914	**.579**	-1.49119	4.66959
	14	1.45780	.68914	**.524**	-1.43594	4.35154
	15	1.44160	.68914	**.531**	-1.45241	4.33561
	16	1.34360	.68914	**.673**	-1.58472	4.27192
3	1	-.45860	.68914	**1.000**	-3.87859	2.96139
	2	2.65340	.68914	**.144**	-.64876	5.95556
	4	-13.20240(*)	.68914	**.000**	-17.13250	-9.27230
	5	-13.39900(*)	.68914	**.000**	-18.33952	-8.45848

	6	-13.20240(*)	.68914	.000	-16.15942	-10.24538
	7	3.89860(*)	.68914	.012	.86980	6.92740
	8	4.25880(*)	.68914	.008	1.18005	7.33755
	9	4.17680(*)	.68914	.022	.56741	7.78619
	10	4.40620(*)	.68914	.019	.72153	8.09087
	11	4.30780(*)	.68914	.008	1.15830	7.45730
	12	4.06220(*)	.68914	.013	1.11035	7.01405
	13	4.24260(*)	.68914	.008	1.15642	7.32878
	14	4.11120(*)	.68914	.009	1.20942	7.01298
	15	4.09500(*)	.68914	.009	1.19288	6.99712
	16	3.99700(*)	.68914	.010	1.06152	6.93248
4	1	12.74380(*)	.68914	.000	8.75021	16.73739
	2	15.85580(*)	.68914	.000	11.92814	19.78346
	3	13.20240(*)	.68914	.000	9.27230	17.13250
	5	-.19660	.68914	1.000	-5.30934	4.91614
	6	.00000	.68914	1.000	-3.79784	3.79784
	7	17.10100(*)	.68914	.000	13.29142	20.91058
	8	17.46120(*)	.68914	.000	13.63590	21.28650
	9	17.37920(*)	.68914	.000	13.26913	21.48927
	10	17.60860(*)	.68914	.000	13.44944	21.76776
	11	17.51020(*)	.68914	.000	13.65691	21.36349
	12	17.26460(*)	.68914	.000	13.33225	21.19695
	13	17.44500(*)	.68914	.000	13.61704	21.27296
	14	17.31360(*)	.68914	.000	13.47754	21.14966
	15	17.29740(*)	.68914	.000	13.45618	21.13862
	16	17.19940(*)	.68914	.000	13.39984	20.99896
5	1	12.94040(*)	.68914	.000	7.98193	17.89887
	2	16.05240(*)	.68914	.000	11.11242	20.99238
	3	13.39900(*)	.68914	.000	8.45848	18.33952
	4	.19660	.68914	1.000	-4.91614	5.30934
	6	.19660	.68914	1.000	-4.78063	5.17383
	7	17.29760(*)	.68914	.000	12.34992	22.24528
	8	17.65780(*)	.68914	.000	12.71992	22.59568
	9	17.57580(*)	.68914	.000	12.57189	22.57971
	10	17.80520(*)	.68914	.000	12.77893	22.83147
	11	17.70680(*)	.68914	.000	12.77463	22.63897
	12	17.46120(*)	.68914	.000	12.28276	22.63964
	13	17.64160(*)	.68914	.000	12.70470	22.57850
	14	17.51020(*)	.68914	.000	12.44209	22.57831
	15	17.49400(*)	.68914	.000	12.41858	22.56942
	16	17.39600(*)	.68914	.000	12.40241	22.38959
6	1	12.74380(*)	.68914	.000	9.62076	15.86684
	2	15.85580(*)	.68914	.000	12.90568	18.80592
	3	13.20240(*)	.68914	.000	10.24538	16.15942

	4	.00000	.68914	1.000	-3.79784	3.79784
	5	-.19660	.68914	1.000	-5.17383	4.78063
	7	17.10100(*)	.68914	.000	14.61260	19.58940
	8	17.46120(*)	.68914	.000	14.87771	20.04469
	9	17.37920(*)	.68914	.000	13.99463	20.76377
	10	17.60860(*)	.68914	.000	14.12424	21.09296
	11	17.51020(*)	.68914	.000	14.80269	20.21771
	12	17.26460(*)	.68914	.000	15.21904	19.31016
	13	17.44500(*)	.68914	.000	14.84798	20.04202
	14	17.31360(*)	.68914	.000	15.20023	19.42697
	15	17.29740(*)	.68914	.000	15.19371	19.40109
	16	17.19940(*)	.68914	.000	14.92666	19.47214
7	1	-4.35720(*)	.68914	.009	-7.53684	-1.17756
	2	-1.24520	.68914	.804	-4.26779	1.77739
	3	-3.89860(*)	.68914	.012	-6.92740	-.86980
	4	-17.10100(*)	.68914	.000	-20.91058	-13.29142
	5	-17.29760(*)	.68914	.000	-22.24528	-12.34992
	6	-17.10100(*)	.68914	.000	-19.58940	-14.61260
	8	.36020	.68914	1.000	-2.33895	3.05935
	9	.27820	.68914	1.000	-3.14283	3.69923
	10	.50760	.68914	1.000	-3.00655	4.02175
	11	.40920	.68914	1.000	-2.39759	3.21599
	12	.16360	.68914	1.000	-2.14216	2.46936
	13	.34400	.68914	1.000	-2.36678	3.05478
	14	.21260	.68914	1.000	-2.11492	2.54012
	15	.19640	.68914	1.000	-2.12512	2.51792
	16	9.8400E-02	.68914	1.000	-2.34434	2.54114
8	1	-4.71740(*)	.68914	.006	-7.93855	-1.49625
	2	-1.60540	.68914	.563	-4.67831	1.46751
	3	-4.25880(*)	.68914	.008	-7.33755	-1.18005
	4	-17.46120(*)	.68914	.000	-21.28650	-13.63590
	5	-17.65780(*)	.68914	.000	-22.59568	-12.71992
	6	-17.46120(*)	.68914	.000	-20.04469	-14.87771
	7	-.36020	.68914	1.000	-3.05935	2.33895
	9	-8.20000E-02	.68914	1.000	-3.53315	3.36915
	10	.14740	.68914	1.000	-3.39303	3.68783
	11	4.9000E-02	.68914	1.000	-2.82259	2.92059
	12	-.19660	.68914	1.000	-2.64498	2.25178
	13	-1.62000E-02	.68914	1.000	-2.79945	2.76705
	14	-.14760	.68914	1.000	-2.59742	2.30222
	15	-.16380	.68914	1.000	-2.60934	2.28174
	16	-.26180	.68914	1.000	-2.80565	2.28205
9	1	-4.63540(*)	.68914	.014	-8.33045	-.94035
	2	-1.52340	.68914	.791	-5.12940	2.08260

	3	-4.17680(*)	.68914	.022	-7.78619	-.56741
	4	-17.37920(*)	.68914	.000	-21.48927	-13.26913
	5	-17.57580(*)	.68914	.000	-22.57971	-12.57189
	6	-17.37920(*)	.68914	.000	-20.76377	-13.99463
	7	-.27820	.68914	1.000	-3.69923	3.14283
	8	8.2000E-02	.68914	1.000	-3.36915	3.53315
	10	.22940	.68914	1.000	-3.67628	4.13508
	11	.13100	.68914	1.000	-3.36642	3.62842
	12	-.11460	.68914	1.000	-3.58152	3.35232
	13	6.5800E-02	.68914	1.000	-3.39004	3.52164
	14	-6.56000E-02	.68914	1.000	-3.45196	3.32076
	15	-8.18000E-02	.68914	1.000	-3.47155	3.30795
	16	-.17980	.68914	1.000	-3.55674	3.19714
10	1	-4.86480(*)	.68914	.011	-8.62941	-1.10019
	2	-1.75280	.68914	.668	-5.43433	1.92873
	3	-4.40620(*)	.68914	.019	-8.09087	-.72153
	4	-17.60860(*)	.68914	.000	-21.76776	-13.44944
	5	-17.80520(*)	.68914	.000	-22.83147	-12.77893
	6	-17.60860(*)	.68914	.000	-21.09296	-14.12424
	7	-.50760	.68914	1.000	-4.02175	3.00655
	8	-.14740	.68914	1.000	-3.68783	3.39303
	9	-.22940	.68914	1.000	-4.13508	3.67628
	11	-9.84000E-02	.68914	1.000	-3.68027	3.48347
	12	-.34400	.68914	1.000	-3.92556	3.23756
	13	-.16360	.68914	1.000	-3.70818	3.38098
	14	-.29500	.68914	1.000	-3.79129	3.20129
	15	-.31120	.68914	1.000	-3.81139	3.18899
	16	-.40920	.68914	1.000	-3.88850	3.07010
11	1	-4.76640(*)	.68914	.006	-8.04802	-1.48478
	2	-1.65440	.68914	.560	-4.79852	1.48972
	3	-4.30780(*)	.68914	.008	-7.45730	-1.15830
	4	-17.51020(*)	.68914	.000	-21.36349	-13.65691
	5	-17.70680(*)	.68914	.000	-22.63897	-12.77463
	6	-17.51020(*)	.68914	.000	-20.21771	-14.80269
	7	-.40920	.68914	1.000	-3.21599	2.39759
	8	-4.90000E-02	.68914	1.000	-2.92059	2.82259
	9	-.13100	.68914	1.000	-3.62842	3.36642
	10	9.8400E-02	.68914	1.000	-3.48347	3.68027
	12	-.24560	.68914	1.000	-2.86944	2.37824
	13	-6.52000E-02	.68914	1.000	-2.94624	2.81584
	14	-.19660	.68914	1.000	-2.80075	2.40755
	15	-.21280	.68914	1.000	-2.81450	2.38890
	16	-.31080	.68914	1.000	-2.98555	2.36395
12	1	-4.52080(*)	.68914	.011	-7.67812	-1.36348

	2	-1.40880	.68914	.505	-4.35193	1.53433
	3	-4.06220(*)	.68914	.013	-7.01405	-1.11035
	4	-17.26460(*)	.68914	.000	-21.19695	-13.33225
	5	-17.46120(*)	.68914	.000	-22.63964	-12.28276
	6	-17.26460(*)	.68914	.000	-19.31016	-15.21904
	7	-.16360	.68914	1.000	-2.46936	2.14216
	8	.19660	.68914	1.000	-2.25178	2.64498
	9	.11460	.68914	1.000	-3.35232	3.58152
	10	.34400	.68914	1.000	-3.23756	3.92556
	11	.24560	.68914	1.000	-2.37824	2.86944
	13	.18040	.68914	1.000	-2.28764	2.64844
	14	4.9000E-02	.68914	1.000	-1.52639	1.62439
	15	3.2800E-02	.68914	1.000	-1.51332	1.57892
	16	-6.52000E-02	.68914	1.000	-2.00429	1.87389
13	1	-4.70120(*)	.68914	.006	-7.92862	-1.47378
	2	-1.58920	.68914	.579	-4.66959	1.49119
	3	-4.24260(*)	.68914	.008	-7.32878	-1.15642
	4	-17.44500(*)	.68914	.000	-21.27296	-13.61704
	5	-17.64160(*)	.68914	.000	-22.57850	-12.70470
	6	-17.44500(*)	.68914	.000	-20.04202	-14.84798
	7	-.34400	.68914	1.000	-3.05478	2.36678
	8	1.6200E-02	.68914	1.000	-2.76705	2.79945
	9	-6.58000E-02	.68914	1.000	-3.52164	3.39004
	10	.16360	.68914	1.000	-3.38098	3.70818
	11	6.5200E-02	.68914	1.000	-2.81584	2.94624
	12	-.18040	.68914	1.000	-2.64844	2.28764
	14	-.13140	.68914	1.000	-2.59831	2.33551
	15	-.14760	.68914	1.000	-2.61045	2.31525
	16	-.24560	.68914	1.000	-2.80377	2.31257
14	1	-4.56980(*)	.68914	.008	-7.66280	-1.47680
	2	-1.45780	.68914	.524	-4.35154	1.43594
	3	-4.11120(*)	.68914	.009	-7.01298	-1.20942
	4	-17.31360(*)	.68914	.000	-21.14966	-13.47754
	5	-17.51020(*)	.68914	.000	-22.57831	-12.44209
	6	-17.31360(*)	.68914	.000	-19.42697	-15.20023
	7	-.21260	.68914	1.000	-2.54012	2.11492
	8	.14760	.68914	1.000	-2.30222	2.59742
	9	6.5600E-02	.68914	1.000	-3.32076	3.45196
	10	.29500	.68914	1.000	-3.20129	3.79129
	11	.19660	.68914	1.000	-2.40755	2.80075
	12	-4.90000E-02	.68914	1.000	-1.62439	1.52639
	13	.13140	.68914	1.000	-2.33551	2.59831
	15	-1.62000E-02	.68914	1.000	-1.76671	1.73431
	16	-.11420	.68914	1.000	-2.14402	1.91562

15	1	-4.55360(*)	.68914	**.008**	-7.64832	-1.45888
	2	-1.44160	.68914	**.531**	-4.33561	1.45241
	3	-4.09500(*)	.68914	**.009**	-6.99712	-1.19288
	4	-17.29740(*)	.68914	**.000**	-21.13862	-13.45618
	5	-17.49400(*)	.68914	**.000**	-22.56942	-12.41858
	6	-17.29740(*)	.68914	**.000**	-19.40109	-15.19371
	7	-.19640	.68914	**1.000**	-2.51792	2.12512
	8	.16380	.68914	**1.000**	-2.28174	2.60934
	9	8.1800E-02	.68914	**1.000**	-3.30795	3.47155
	10	.31120	.68914	**1.000**	-3.18899	3.81139
	11	.21280	.68914	**1.000**	-2.38890	2.81450
	12	-3.28000E-02	.68914	**1.000**	-1.57892	1.51332
	13	.14760	.68914	**1.000**	-2.31525	2.61045
	14	1.6200E-02	.68914	**1.000**	-1.73431	1.76671
	16	-9.80000E-02	.68914	**1.000**	-2.11638	1.92038
16	1	-4.45560(*)	.68914	**.007**	-7.56316	-1.34804
	2	-1.34360	.68914	**.673**	-4.27192	1.58472
	3	-3.99700(*)	.68914	**.010**	-6.93248	-1.06152
	4	-17.19940(*)	.68914	**.000**	-20.99896	-13.39984
	5	-17.39600(*)	.68914	**.000**	-22.38959	-12.40241
	6	-17.19940(*)	.68914	**.000**	-19.47214	-14.92666
	7	-9.84000E-02	.68914	**1.000**	-2.54114	2.34434
	8	.26180	.68914	**1.000**	-2.28205	2.80565
	9	.17980	.68914	**1.000**	-3.19714	3.55674
	10	.40920	.68914	**1.000**	-3.07010	3.88850
	11	.31080	.68914	**1.000**	-2.36395	2.98555
	12	6.5200E-02	.68914	**1.000**	-1.87389	2.00429
	13	.24560	.68914	**1.000**	-2.31257	2.80377
	14	.11420	.68914	**1.000**	-1.91562	2.14402
	15	9.8000E-02	.68914	**1.000**	-1.92038	2.11638

The mean difference is significant at the .05 level.
Factor 1 – 16 refers to model 1 – 16 separately

Appendix III ANOVA Results of 3 Methods' Prediction Accuracies

Multiple Comparisons
Dependent Variable: TOTAL
Games-Howell

(I) FACTOR2	(J) FACTOR2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
10	20	-13.99953(*)	.42719	.000	-15.43366	-12.56540
	30	3.42402(*)	.34441	.000	2.20171	4.64633
20	10	13.99953(*)	.42719	.000	12.56540	15.43366
	30	17.42355(*)	.34441	.000	16.47251	18.37460
30	10	-3.42402(*)	.34441	.000	-4.64633	-2.20171
	20	-17.42355(*)	.34441	.000	-18.37460	-16.47251

The mean difference is significant at the .05 level.
Factor 10 refers to LRA methods
Factor 20 refers to TRI methods
Factor 30 refers to ANN methods

Appendix IV ANOVA Results of 7 Models' Time Consumption

(I) FACTOR1	(J) FACTOR1	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-.20640	.59835	1.000	-3.48419	3.07139
	3	.14700	.59835	1.000	-3.09663	3.39063
	4	1.43360	.59835	.533	-1.78568	4.65288
	5	1.65360	.59835	.400	-1.57584	4.88304
	6	1.75020	.59835	.371	-1.45916	4.95956
	7	-.23620	.59835	1.000	-3.44844	2.97604
2	1	.20640	.59835	1.000	-3.07139	3.48419
	3	.35340	.59835	.997	-2.10305	2.80985
	4	1.64000	.59835	.173	-.59447	3.87447
	5	1.86000	.59835	.105	-.36221	4.08221
	6	1.95660	.59835	.097	-.31467	4.22787
	7	-2.98000E-02	.59835	1.000	-2.28290	2.22330
3	1	-.14700	.59835	1.000	-3.39063	3.09663
	2	-.35340	.59835	.997	-2.80985	2.10305
	4	1.28660	.59835	.302	-.79690	3.37010
	5	1.50660	.59835	.176	-.55813	3.57133
	6	1.60320	.59835	.167	-.52965	3.73605
	7	-.38320	.59835	.987	-2.49236	1.72596
4	1	-1.43360	.59835	.533	-4.65288	1.78568
	2	-1.64000	.59835	.173	-3.87447	.59447
	3	-1.28660	.59835	.302	-3.37010	.79690
	5	.22000	.59835	.997	-1.31771	1.75771
	6	.31660	.59835	.988	-1.38899	2.02219
	7	-1.66980(*)	.59835	.048	-3.32420	-1.54026E-02
5	1	-1.65360	.59835	.400	-4.88304	1.57584
	2	-1.86000	.59835	.105	-4.08221	.36221
	3	-1.50660	.59835	.176	-3.57133	.55813
	4	-.22000	.59835	.997	-1.75771	1.31771
	6	9.6600E-02	.59835	1.000	-1.56588	1.75908
	7	-1.88980(*)	.59835	.021	-3.49634	-.28326
6	1	-1.75020	.59835	.371	-4.95956	1.45916

	2	-1.95660	.59835	.097	-4.22787	.31467
	3	-1.60320	.59835	.167	-3.73605	.52965
	4	-.31660	.59835	.988	-2.02219	1.38899
	5	-9.66000E-02	.59835	1.000	-1.75908	1.56588
	7	-1.98640(*)	.59835	.026	-3.74295	-.22985
7	1	.23620	.59835	1.000	-2.97604	3.44844
	2	2.9800E-02	.59835	1.000	-2.22330	2.28290
	3	.38320	.59835	.987	-1.72596	2.49236
	4	1.66980(*)	.59835	.048	1.5403E-02	3.32420
	5	1.88980(*)	.59835	.021	.28326	3.49634
	6	1.98640(*)	.59835	.026	.22985	3.74295

The mean difference is significant at the .05 level.
Factor 1 – 6 refers to model 1 – 6, factor 7 refers to ANN model with 1 layer 3 hidden unit

Appendix V Variable Definition Table

Variable	Definition
ofr	Number of Offers sent
ofrwmsg	Number of Offers with messages sent
msg	Number of messages sent
agr	Status of Agreement
escore	Expected Score
rscore	Reservation score
score	Actual Sores
opt	Was agreement optimal?
ps_rchd	Did system suggest ps mechanism?
ps_used	If suggested, was the ps mechanism used?
ps_ofr	Number of offers in ps
nego_len	Length of Negotiations
opt_scr	Optimal score
dlineexp	Did deadline expire
conf_use	(Score < Rscore) is true
mnofr_ln	Mean time between offers in seconds
mnmsg_ln	Mean length of messages
act_dln	Activity 48 hrs before deadline
off_1	No. of ofrs 1 day before deadline
off_2	No. of ofrs 2 day before deadline
off_4	No. of ofrs 4 day before deadline
off_all	No. of ofrs sent at least 5 days before deadline
time_ds	Time dist. between agr and deadline
caseund	Ease to understand case
cborn	Country of birth
creside	Country of residence
efrndly	Expected friendliness of negotiation
imore	Increase Internet access (pre)
yofb	Year of birth
nexp	Negotiation experience
iacc	Present internet access
agrsat	Satisfaction with agreement
control	Level of perceived control
mete	Met expectations
frndly	Friendliness of negotiation
perf	Personal performance
after	Internet access(in future)
gender	User's gender
firstlan	First Language
knewoppc	Know opp country (pre)
knewoppi	Know opponent's ID
nssbefor	Used NSS Before
wtgissue	Ease of weighing issues

wtgoptio	Ease of weighing options
difforig	If ps used, did agreement change
graphuse	Was Graph Used
graphinf	Was Graph informative
graphme	Did Graph Influence Me
graphas	Did Graph Influence Assessment
instruct	Ease of system instructions
ineasy	InterNeg Easy
knewopp	Knew Opponent(post)
msghelpf	Msg Helpful
practice	Use Ineg for practice
prepare	Use Ineg for preparation
nego	Use Ineg for negos
oppcoop	Opponent co-operative
cguess	Guess opponent's country
disclc	Did opponent disclose country
disclid	Did opponent disclose ID
oppexplo	OppExploitative
opphones	Opp Honest
oppinfor	Opp Informative
opppersu	Opp Persuasive
psusedq	Use post-settlement (user-entered)
predicto	Learning to Predict
seeopp	Interested to see opp
settleme	Accept same settlement for real
surprise	Surprised by opp
understo	Understood Opp Priorities
utilityv	Importance of Utility Values
workwopp	Work with opp
langrp	Language Groups
langnum	Language groups numbered
filter_$	opt_scr > 79.9 (FILTER)
age	Age of users (computed in 1998)
ofr_befr	Offers before post-settlement
dif_scor	Difference between achieved and expected scores

VI. Transformation of Selected Interval Variables

Name	Keep	Role	Formula	Skew	Label
AGE	No	input		2.3792093248	AGE
AGE_QAZZ	Yes	input	AGE	2.3792093248	Bucket(AGE)
ESCORE	No	input		-3.809551134	ESCORE
ESCO_AR7	Yes	input	((ESCORE + 281))** 2	-1.348312877	ESCORE: Maximize normality
LANGNUM	No	input		0.9530091719	LANGNUM
LANG_ZZL	Yes	input	(LANGNUM)** 0.25	0.3219799759	LANGNUM: Maximize normality
MNMSG_LN	No	input		12.51533059	MNMSG_LN
MNMS_I92	Yes	input	(MNMSG_LN)** 0.25	0.608852576	MNMSG_LN: Maximize normality
MNOFR_LN	No	input		1.9057335927	MNOFR_LN
MNOF_MBM	Yes	input	sqrt((MNOFR_LN + 1))	0.1919247387	MNOFR_LN: Maximize normality
MSG	No	input		2.546660011	MSG
MSG_9JCB	Yes	input	log((MSG + 1))	0.3611657077	MSG: Maximize normality
NEGO_LEN	No	input		1.3245523431	NEGO_LEN
NEGO_G9T	Yes	input	sqrt((NEGO_LEN + 15513))	0.664185429	NEGO_LEN: Maximize normality
OFF_ALL	No	input		1.113771362	OFF_ALL
OFF__8E7	Yes	input	((OFF_ALL + 1))** 0.25	-0.084666597	OFF_ALL: Maximize normality
OFR	No	input		0.8668532432	OFR
OFR_N3FX	Yes	input	sqrt((OFR + 1))	0.068279131	OFR: Maximize normality
OFRWMSG	No	input		0.581028195	OFRWMSG
OFRW_J74	Yes	input	sqrt((OFRWMSG + 1))	-0.017833364	OFRWMSG: Maximize normality
OPT_SCR	No	input		-3.189466076	OPT_SCR
OPT__6NQ	Yes	input	exp(OPT_SCR)	-0.741114062	OPT_SCR: Maximize normality
RSCORE	No	input		-3.226993852	RSCORE
RSCO_HXH	Yes	input	((RSCORE + 301))**	-0.967964045	RSCORE:

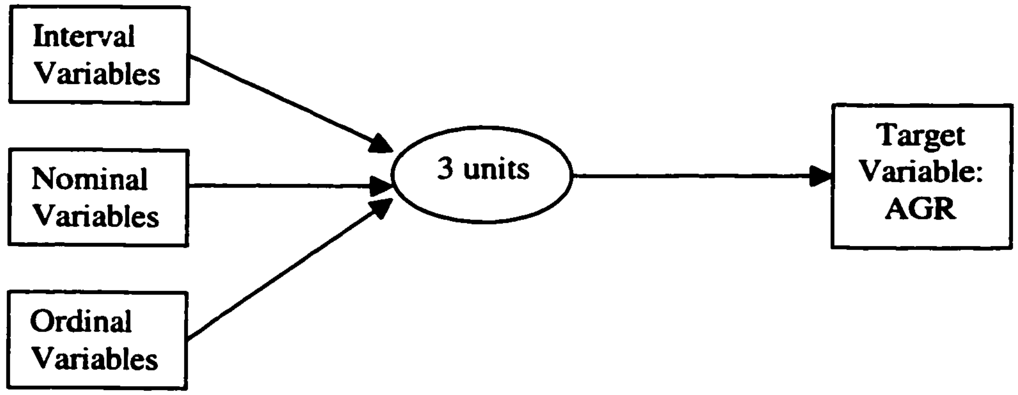
			2		Maximize normality
DIF_SCOR	Yes	rejected		-0.855612255	DIF_SCOR
OFR_BEFR	Yes	rejected		1.3986537289	OFR_BEFR
SCORE	Yes	rejected		-3.020674118	SCORE
TIME_DS	Yes	rejected		8.8600811171	TIME_DS
YOFB	Yes	rejected		-2.379209337	YOFB

* The absolute value of Skew indicates the skewness. The bigger it is, the more skew it is. A 0 indicates perfectly normally distributed.
** The transformed variable is used for data mining and the original one is not. A transformed variable is listed exactly under the original one. For example, the variable ESCORE (skew is –3.809) is transformed to a new variable, named ESCORE_AR7 (skew is –1.348); ESCORE_AR7 is not used for data mining instead of ESCORE. The transformation formula is ESCORE_AR7 = ((ESCORE + 281))** 2.

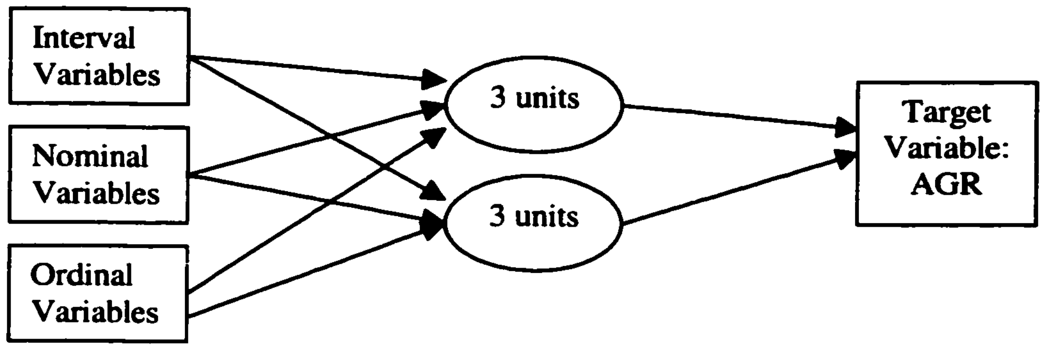
Appendix VII ANN Models:

Group1

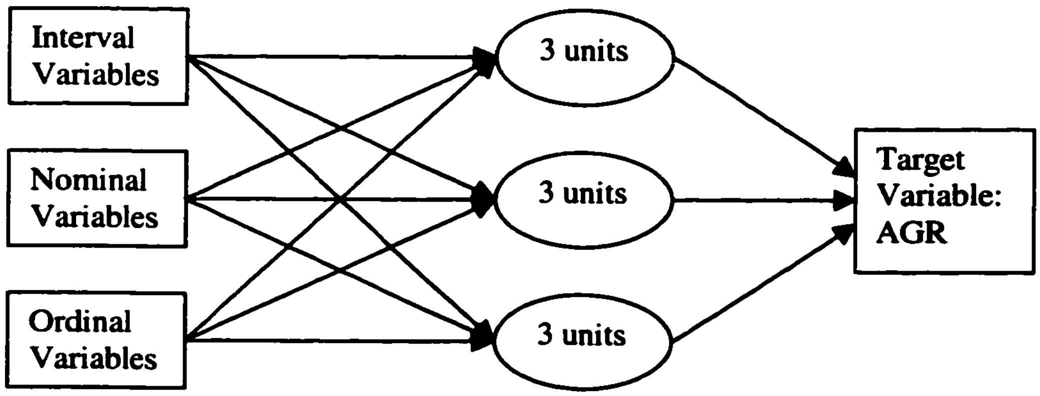
Model 7: 1 layer, 3 hidden units



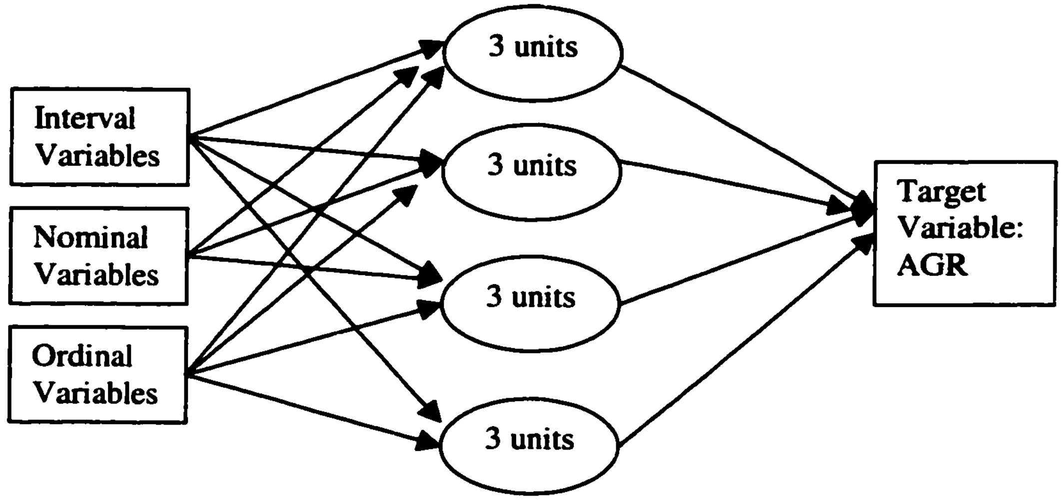
Model 8.: 2 layers, 3 hidden units



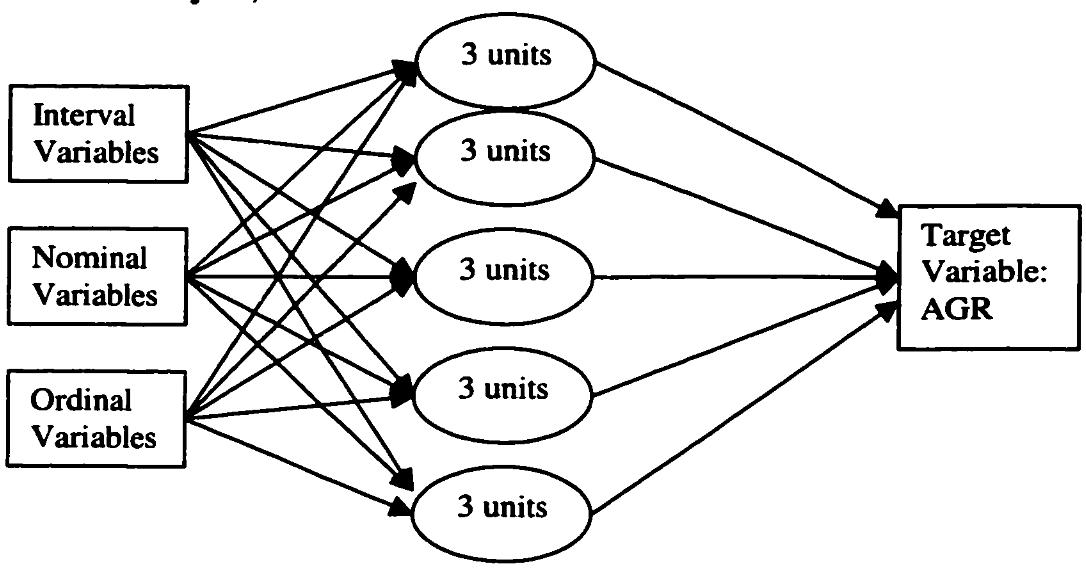
Model 9: 3 layers, 3 hidden units



Model 10: 4 layers, 3 hidden units

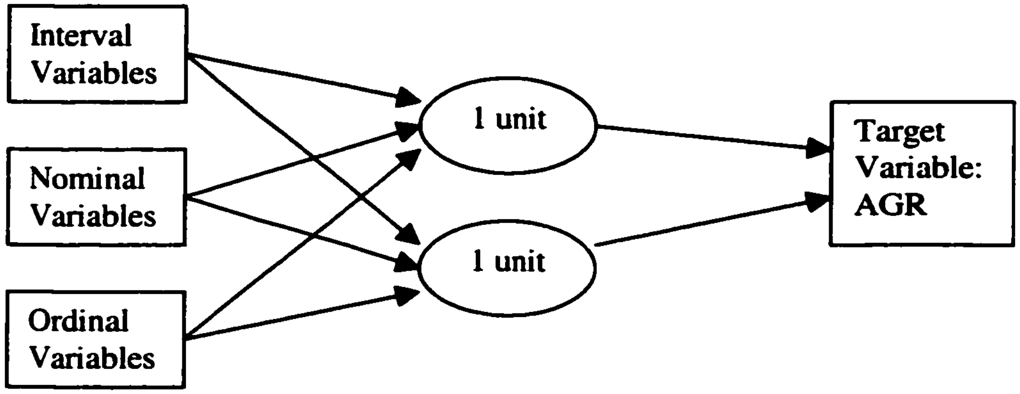


Model 11: 5 layers, 3 hidden units

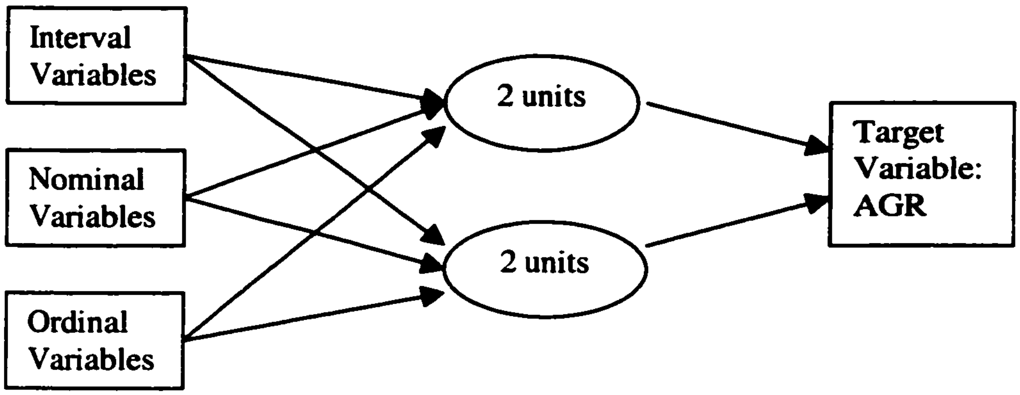


Group 2

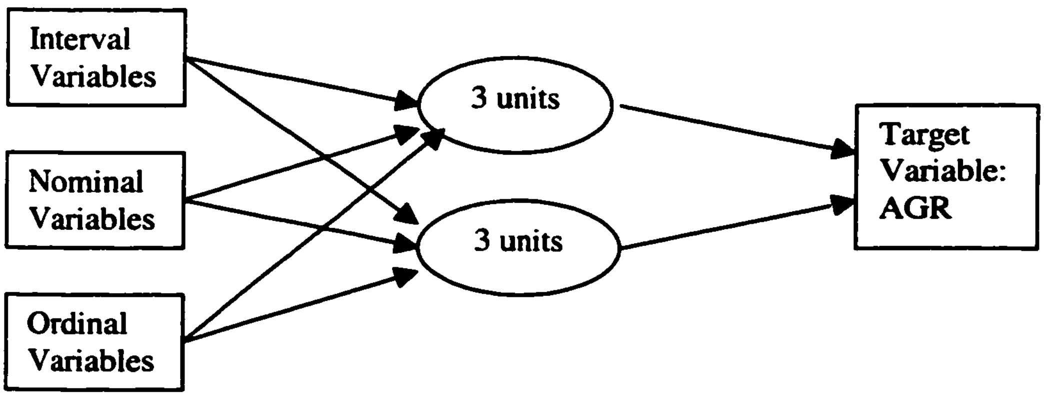
Model 12: 2 layers, 1 hidden unit



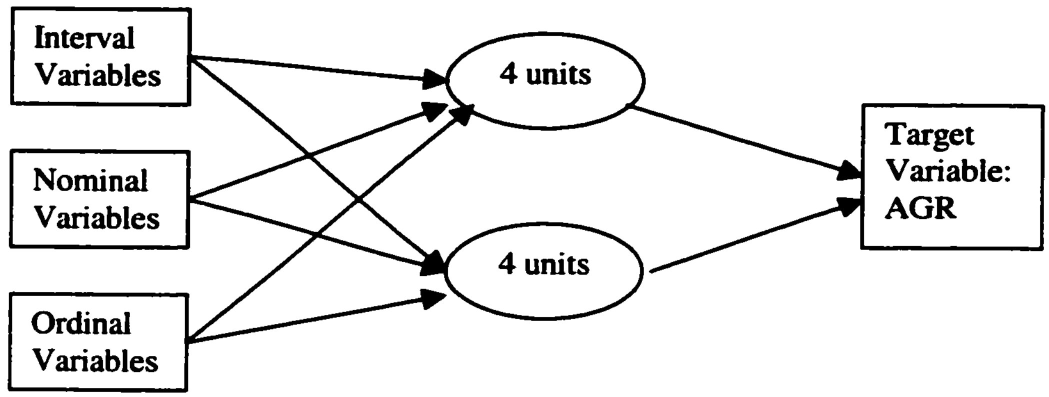
Model 13: 2 layers, 2 hidden units



Model 14: 2 layers, 3 hidden units



Model 15: 2 layers, 4 hidden units



Model 15: 2 layers, 5 hidden units