DEANFIS – A FUZZY MULTI-DIMENSIONAL APPROACH TO DECISION-MAKING USING DATA ENVELOPMENT ANALYSIS AND NEURO-FUZZY SYSTEMS

Rashmi Malhotra, St. Joseph's University, (610) 660-3497, rmalhotr@sju.edu

ABSTRACT

To screen consumer loan applications, loan officers use many different methods besides intuitive judgment and experience. They use mathematical techniques such as credit-scoring models and traditional statistical models. In addition, many financial institutions use artificial intelligence methods such as expert systems, artificial neural systems, and fuzzy logic. This study proposes the development of a decision support system that uses a combination of data envelopment analysis and neuro-fuzzy systems. Thus, the decision support system derives benefit from both methodologies to provide a comprehensive review of a loan applicant.

Keywords: Benchmarking, Data Envelopment Analysis, Neuro-fuzzy systems, Decision Support Systems.

INTRODUCTION

A business organization's objective is to make better decisions at all levels of the firm to improve performance. Typically organizations are multi-faceted and complex systems that use uncertain information. Therefore, making quality decisions to improve organizational performance is a daunting task. Organizations use decision support systems that apply different business intelligence techniques such as statistical models, scoring models, neural networks, expert systems, neuro-fuzzy systems, case-based systems, or simply rules that have been developed through experience. Managers need a decision-making approach that is robust, competent, effective, efficient, and integrative to handle the multi-dimensional organizational entities. The decision maker deals with multiple players in an organization such as products, customers, competitors, location, geographic structure, scope, internal organization, and cultural dimension (Porter, 1980). Sound decisions include two important concepts: efficiency (return on invested resources) and effectiveness (reaching predetermined goals). However, quite frequently, the decision maker cannot simultaneously handle data from different sources. Hence, I recommend that managers analyze different aspects of data from multiple sources separately and integrate the results of the analysis. This study proposes the design of a multi-attribute-decision-support-system that combines the analytical power of two different tools: data envelopment analysis (DEA) and fuzzy logic. DEA evaluates and measures the relative efficiency of decision making units that use multiple inputs and outputs to provide non-objective measures without making any specific assumptions about data. On the other hand fuzzy logic's main strength lies in handling imprecise data. This study proposes a modeling technique that jointly uses the two techniques to benefit from the two methodologies. A major advantage of the DEA approach is that it clearly identifies the important factors contributing to the success of a decision. In addition, I also propose the use of a neuro-fuzzy model to create a rule-based system that can aid the decision-maker in making decisions regarding the implications of a decision. One of the important characteristics of neuro-fuzzy systems is their ability to deal with imprecise and uncertain information. The neuro-fuzzy model integrates the performance values of a set of production units derived by ranking using DEA to create IF-THEN rules to handle fluctuating and uncertain scenarios. Thus, a decision maker can easily analyze and understand any decision made by the neuro-fuzzy model in the form of the easily interpretable IF-THEN rules. Finally, this study will extend my earlier work on the application of neural network models and neuro-fuzzy models to decision-making. In these studies, I compared the performance of neural network systems and neuro-fuzzy systems with statistical models. In both cases, neural systems and neuro-fuzzy systems outperformed the statistical models. I propose to further improve the neuro-fuzzy models by combining them with DEA models.

LITERATURE REVIEW

Neural Networks and Neuro-Fuzzy Literature

Many studies highlight the use of artificial neural systems in business applications. Anders, Korn, and Schmitt (1998) use statistical inference techniques to build neural network models to explain the prices of call options on the German stock index DAX. They show that statistical specification strategies lead to parsimonious networks that have a superior out-of-sample performance when compared to the Black-Scholes model. Ntungo and Boyd (1998) report that out-of-sample neural network trading returns for corn, silver, and Deutsche mark futures contracts are positive and at about the levels as the returns with ARIMA models. Desai and Bharati (1998) test the efficacy of neural networks in predicting returns on stock and bond indices. They find that the neural network forecasts are conditionally efficient with respect to linear regression models for large stocks and corporate bonds, whereas the evidence is not statistically significant for small stocks and intermediate-term government bonds.

Zhang and Hu (1998) illustrate the use of neural networks in forecasting UK pound/U.S. dollar exchange rate. They report that neural networks outperform linear models, particularly when the forecast horizon is short. Zhang, Hu, Patuwo, and Indro (1999) show that neural networks are significantly outperform logistics regression models in bankruptcy prediction. Indro, Jiang, Patuwo, and Zhang (1999) show that neural networks outperform linear models in forecasting the performance of mutual funds that follow value, blend, and growth investment styles. Thus, all the above-mentioned studies provide mixed evidence regarding the potential of

neural networks to analyze, evaluate, and predict many financial systems. Wong, Wang, Goh, and Quek (1992) illustrate the use of fuzzy neural systems for stock selection. Derrig and Ostaszewski (1995) provide an overview of fuzzy pattern recognition techniques and used those techniques in clustering for risk and claims classification. Turtle and Gill (1994) illustrate the use of fuzzy logic in international cash management. Selwyn (1999) compares the benefits of using neurofuzzy systems and neural networks for credit-risk evaluation. Malhotra and Malhotra (2002) use neurofuzzy systems to evaluate consumer units/decisions. They find that the neuro-fuzzy system performs better than the multiple discriminant analysis approach to identify bad credit applications. In addition, Malhotra and Malhotra (2003) also find that the neural network models consistently perform better than the multiple discriminant analysis models in identifying potential problem units/decisions. Angelini et. al. (2008) show that artificial neural networks show promising results in evaluating credit risk. Wu (2008) illustrates several data mining algorithms, i.e., Backpropagation Neural Network (BPNN), decision tree, logistic regression and Monte Carlo simulation in credit-scoring problem. Hu et. al. (2007) illustrate the use of neural network to evaluate credit risk in relation to the retail industry. Trinkle et. al. (2007) illustrate the use of neural networks in developing credit risk models. Baesens et.al. (2003) illustrate the use of neural networks for extracting rules and creating decision tables for credit risk evaluation. Charlton, et. al. (2010) illustrate the use of neuro-fuzzy systems to identify chemical parameters including isotopic ratios, biogenic amines, and rare earth elements that discriminate between vintages and grape varieties for wines produced in a particular country. Fernandes, et. al. (2010) illustrate the use of Adaptive Network-based Fuzzy Inference Systems (ANFIS) that classifies the mammographic images of calcification region of interest as benign or malign. Bekiros (2010) use hybrid neurofuzzy system for decision-making and trading under uncertainty. Pinthong, et. al. (2009) illustrate the use of hybrid genetic algorithm and neurofuzzy computing to determine the optimal reservoir releases in Pasak River Basin, Thailand. Zhou et. al. (2004) use neurofuzzy approach to discover knowledge for predicting customers' attitude toward Internet retailers in relation to traditional retailers. Yang et. al. (2002) illustrate the use of neuro-fuzzy systems in early detection and assessment of a surface mount assembly problem for cost-effective manufacturing. This study extends previous studies by applying DEA models and neurofuzzy systems to analyze consumer units/decisions.

Data Envelopment Analysis Literature

Recently, many studies have illustrated the use of DEA, a non-parametric methodology to analyze different aspects of business entities. The details of the DEA model are discussed in the next section. In contrast to other methodologies, DEA is one of the methods that have traditionally been used to assess the comparative efficiency of homogenous operating units such as schools, hospitals, utility companies, sales outlets, prisons, and military operations. More recently, it has been applied to banks (Haslem, Scheraga, & Bedingfield, 1999), railroad efficiency, airport efficiency, retailing, and mutual funds (Haslem & Scheraga, 2003; Galagedera & Silvapulle, 2002; McMullen & Strong, 1998; Murthi, Choi, & Desai, 1997). Zhu (2000) uses data envelopment analysis to develop a multi-factor financial performance model that recognizes tradeoffs among various financial measures. Kao and Liu (2004) compute efficiency scores based on the data contained in the financial statements of Taiwanese banks. They use this data to make advanced predictions of the performances of 24 commercial banks in Taiwan. Pille and Paradi (2002) analyze the financial performance of Ontario credit unions. They develop models to detect weaknesses in Credit Unions in Ontario, Canada. Yasar and McCure (1996) use data envelopment analysis for measuring and assessing the financial performance for hospitals. They compute a financial performance index (FPI) as a measure of aggregate financial performance. They show that financial performance index across many financial ratios eases the comparison of an individual hospital with its peers. Halkos and Salamouis (2004) explore the efficiency of Greek banks with the use of a number of suggested financial efficiency ratios for the time period 1997-1999. They show that data envelopment analysis can be used as either an alternative or complement to ratio analysis for the evaluation of an organization's performance. The study finds that the higher the size of total assets the higher the efficiency. Neal (2004) investigates X-efficiency and productivity change in Australian banking between 1995 and 1999 using data envelopment analysis and Malmquist productivity indexes. It differs from earlier studies by examining efficiency by bank type, and finds that regional banks are less efficient than other bank types. The study concludes that diseconomies of scale set in very early, and hence are not a sufficient basis on which to allow mergers between large banks to proceed. Paradi and Schaffnit (2004) evaluate the performance of the commercial branches of a large Canadian bank using data envelopment analysis. Chen, Sun, and Peng (2005) study the efficiency and productivity growth of commercial banks in Taiwan before and after financial holding corporations' establishment. They employ a data envelopment analysis approach to generate efficiency indices as well as Malmquist productivity growth indices for each bank. Howland and Rowse (2006) assess the efficiency of branches of a major Canadian bank by benchmarking them against the DEA model of American bank branch efficiency. Sufian (2007) uses DEA approach to evaluate trends in the efficiency of the Singapore banking sector. The paper uses DEA approach to distinguish between technical, pure technical and scale efficiencies.

Sanjeev (2007) evaluates the efficiency of the public sector banks operating in India for a period of five years (1997-2001) using DEA. The study also investigates if there is any relationship between the efficiency and size of the banks. The results of the study suggest that no conclusive relationship can be established between the efficiency and size of the banks. Lin, Shu, and Hsiao (2007) study the relative efficiency of management in the Taiwanese banking system through DEA. The goal is to estimate the competitiveness of each bank and managerial efficiency is to show the efficiency variation of

each bank through Malmquist index. Bergendahl and Lindblom (2008) develop principles for an evaluation of the efficiency of a savings bank using data envelopment analysis as a method to consider the service orientation of savings banks. They determine the number of Swedish savings banks being "service efficient" as well as the average degree of service efficiency in this industry. Hoon and Chunyan (1994) analyzed the productive efficiency of the railway services in 19 Organization for Economic Cooperation and Development (OECD) countries. They report that railway systems with high dependence on public subsidies are less efficient than similar railways with less dependence on subsidies. Cowie and Riddington (1996) evaluate the efficiency of the European railways through the use of a production frontier approach. Yu and Lin (2008) uses a multi-activity network DEA model to simultaneously estimate passenger and freight technical efficiency, service effectiveness, and technical effectiveness for 20 selected railways for the year 2002. Lozano & Gutierrez (2011) illustrate the slacks-based measure of efficiency of 39 Spanish airports using DEA. Liu & Liu (2010) illustrate the use of DEA in evaluating and ranking the research and redevelopment performance of Taiwan's government-supported research institutes. Saranga & Moser (2010) develop a comprehensive performance measurement framework using the classical and two-stage Value Chain Data Envelopment Analysis model.

Sellers-Rubio and Mas-Ruiz (2006) examined the efficiency of supermarket chains in the Spanish retailing industry using DEA. The study evaluated the 100 supermarket chains between 1995 and 2001 revealing that there were "high levels of economic inefficiency in the Spanish retailing industry" (Sellers-Rubio and Mas-Ruiz, 2006). The study found that the "underlying causes of the difference between the current performance level of an organization and the best practically possible level are management style, organizational structure, and product quality among others" (Sellers-Rubio and Mas-Ruiz, 2006). Donthu and Yoo (1998) "measured the relative-to-best performance efficiency of retail outlets characterized by multiple inputs and outputs using data collected from retail stores belonging to a restaurant chain". The greatest advantage of using DEA to evaluate performance efficiency is that a retail outlet is compared to the best performing retail outlets, otherwise known as benchmarking. Through benchmarking, retail outlets of a chain store or a franchise system can be comparatively evaluated for performance, ultimately improving the operation of the entire retail store. Jiang and Talaga (2006) used DEA to explore the relationship between satisfying customers and building a customer base for the e-tailing industry. The study found that "performance scores for developing a customer base vary across product categories" (Pingjun and Talaga, 2006), and "performance score is a good parameter for predicting future change on a unique number of visitors and on the competition pattern for a particular e-tailer" (Pingjun and Talaga, 2006). This study used DEA to help e-tailers "measure the reach efficiency which is the extent to which the e-tailer is attracting visitors within an online environment and page-view efficiency measures the performance of an e-store in generating more page view per visitor" (Pingjun and Talaga, 2006). By simultaneously getting more visitors to an e-tailer's web site and having them stay longer is reflected in the overall efficiency of the unit (Pingjun and Talaga, 2006). Barth (2007) used DEA to show that "new-style retail wine stores with features such as tasting rooms, lecture theatres, and demonstration kitchens used to educate and engage customers have better retail efficiency than old-style stores". The DEA used "sales dollars, labour hours, and litres of inventory depletion from a paired-sample of old-style and new-style facilities to determine the retail efficiency of the stores" (Barth, 2007). The results of the study reflected that the new-style stores had higher retail efficiency than the old style stores and reducing the input in the older stores does not increase the retail efficiency of these stores. Although the study shows that the retail efficiency is increased with the new store features, the contribution of each feature towards the overall improvement in retail performance is unknown (Barth, 2007). Athanassopoulos and Thanassoulis (1995) used DEA to assess the market effieciency of pubs in the UK to aid in planning. Market efficiency is defined as "the extent to which a unit is exploiting the potential within its catchment area for generating revenue (Athanassopoulos and Thanassoulis, 1995). The inputs are environmental variables as well as one uncontrollable internal variable and the output is the revenue generated (Athanassopoulos and Thanassoulis, 1995). Pilling, et al (1999) used to adjust salesperson performance for territory characteristics in order to give a more individualized and complete picture of salesperson performance. DEA helped adjust for territory inequities in the evaluation process and identified best practices among a group of salespeople as well as included ways to increase the impact of sales-force related expenditures (Pilling, et al, 1999).

Murthi, Choi, & Desai (1997) examine the market efficiency of the mutual fund industry by different investment objectives. They use a benefit/cost non-parametric analysis where a relationship between return (benefit) and expense ratio, turnover, risk, and loads (cost) is established. They also develop a measure of performance of mutual funds that has a number of advantages over traditional indices. The DEA portfolio efficiency index (DEPI) does not require specification of a benchmark, but incorporates transaction costs. The most important advantage of DEA method as compared to other measures of fund performance is that DEA identifies the variables leading to inefficiencies and the levels by which they should be changed to restore the fund to its optimum level of efficiency. McMullen and Strong (1998) applied DEA to evaluate the relative performance of 135 US common stock funds using one, three, and five-year annualized returns, standard deviation of returns, sales charge, minimum initial investment, and expense ratio. They illustrate that DEA can assist in selecting mutual funds for an investor with a multifactor utility function. The DEA selects optimum combinations of investment characteristics, even when the desired characteristics are other than the two-factors specified in Capital Market Theory. The DEA enable the user to determine the most desirable alternatives, and pinpoint the inefficiencies in a DEA-inefficient alternative. Sedzro and Sardano (1999) analyzed 58 US equity funds in Canada using DEA with annual return, expense ratio, minimum initial

investment and a proxy for risk as factors associated with fund performance. Further, they also find a strong relationship among the efficiency rankings using DEA, Sharpe ratios, and Morningstar data. Galagedera and Silvapulle (2002) use DEA to measure the relative efficiency of 257 Australian mutual funds. The further investigate the sensitivity of DEA efficiency to various input-output variable combinations. They find that more funds are efficient when DEA captures a fund's long-term growth and income distribution than a shorter time horizon. In general, the overall technical efficiency and the scale efficiency are higher for risk-aversive funds with high positive net flow of assets.

Haslem and Scheraga (2003) use DEA to identify efficiencies in the large-cap mutual funds in the 1999 Morningstar 500. They identify the financial variables that differ significantly between efficient and inefficient funds, and determine the nature of the relationships. They use Sharpe index as the DEA output variable. They find that the input/output and profile variables are significantly different between the Morningstar 500 (1999) large-cap mutual funds that are DEA performance-efficient and inefficient. Basso and Funari (2001) propose the use of DEA methodology to evaluate the performance of mutual funds. The proposed DEA performance indexes for mutual funds represent a generalization of various traditional numerical indexes that can take into account several inputs and outputs. They propose two classes of DEA indexes. The first class generalizes the traditional measures of evaluation using different risk indicators and subscription and redemption costs that burden the fund investment. The second class of indexes considers a multiple inputs-outputs structure. Thus, they monitor not only the mean return but also other features such as stochastic dominance and the time lay-out. Morey and Morey (1999) present two basic quadratic programming approaches for identifying those funds that are strictly dominated, regardless of the weightings on different time horizons being considered, relative to their mean returns and risks. They present a novel application of the philosophy of data envelopment analysis that focuses on estimating "radial" contraction/expansion potentials. Furthermore, in contrast to many studies of mutual fund's performance, their approach endogenously determines a custom-tailored benchmark portfolio to which each mutual fund's performance is compared. Feroz, Kim, and Raad (2003) illustrate the use of data envelopment analysis to evaluate the financial performance of oil and gas industry. Edirisinghe and Zhang (2007) develop a data envelopment analysis model to evaluate a firm's financial statements over time in order to determine a relative financial strength indicator that can predict firm's stock price returns.

Studies combining DEA and Neuro-fuzzy models:

Recently, some studies have proposed the fusion or joint modeling of DEA models and fuzzy logic. Omero, et. al. (2005) illustrate the development of a decision support system that uses DEA for qualitative data analysis. The system uses fuzzy logic to integrate heterogeneous data from multiple sources. Triantis (2003) propose a fuzzy DEA approach that computes fuzzy non-radial technical efficiency measures and implements the approach for a newspaper preprint insertion manufacturing process. Wu (2009) proposes an integrated approach to rate decision alternatives using data envelopment analysis and preference relations in three stages. First, pair wise efficiency scores are computed using two DEA models: the CCR model and the proposed cross-evaluation DEA model. Second, the pair wise efficiency scores are then utilized to construct the fuzzy preference relation and the consistent fuzzy-preference relation. Third, by use of the row wise summation technique, the study yields a priority vector that is used to rank decision-making units. Zeydan et. al. (2009) illustrate a new framework that combines fuzzy TOPSIS (technique for order preference by similarity to ideal solution to measure qualitative performance and DEA to measure quantitative performance. Hougaard (1999) suggests extending technical efficiency scores of DEA models to fuzzy scores that enable the decision maker to use scores of technical efficiency in combination with other sources of available performance e.g. expert opinions, key figures, etc. In addition, many researchers have developed models to combine fuzzy logic and DEA to handle imprecise or vague data. Sengupta (1992) utilized the probabilistic feasibility of the inequality constraints to propose a fuzzy approach and use a fuzzy linear programming transformation approach as a viable approach. Many researchers have proposed fuzzy mathematical programming approaches such as probabilistic programming and alpha-cut approaches to assess the relative efficiency of the DMUs (Guo & Tanaka, 2001; Lertworasirikul, Fang, Joines, & Nuttle, 2003; Leon, Liern, Ruiz, & Sirvent, 2003; Saati, Memariani & Jahanshahahloo, 2002). Lertworasirikul et. al. (2003) propose a possibility approach to the treatment of fuzzy DEA models. Guo & Tanaka (2001) introduce an alpha-cut approach that changes a fuzzy DEA model to a bi-level LP model. Kao & Liu (2000) propose a technique to transform a fuzzy DEA model into a family of crisp DEA models by applying the alpha-cut approach. They solve multiple LP problems to approximate the membership function of the efficiency score to assess a DMU. Liu (2008) developed a fuzzy DEA model to find the efficiency measures embedded with the assurance region concept. Shokouhi et. al. (2010) propose a robust optimization method to deal with data uncertainties that cover the interval approach DEA results with fuzzy DEA approach. Ma & Li (2008) propose a methodology to incorporate fuzzy preferences and range reduction techniques. The study first adopts a modified DEA model to generate reasonable upper and lower bounds of preference ratios. By referring to these ranges, a decision maker then specifies his/her preferences partially. Qin & Liu (2010) present several formulas for mean chance distributions of triangular fuzzy random variables and their functions to develop a new class of fuzzy random data envelopment analysis models. Kao & Liu (2003) use maximizing and minimizing set methods to rank the fuzzy efficiency scores without knowing the exact form of the membership function.

As illustrated above, all of these studies illustrate the merger of fuzzy logic and DEA to develop fuzzy DEA models. None of the studies illustrate the use of DEA models and fuzzy

logic to form DEANFIS model, in this order. Therefore, the purpose of this study is twofold. Firstly, this study investigates and analyzes the synergy of DEA and ANFIS models theoretically to develop DEANFIS models. The study further plans to use simulated data to validate the models. Secondly, this study aims to use industry data to design decision support system that can provide rules to the manager to make decisions.

METHODOLOGY

The Data Envelopment Analysis Model

The Data Envelopment Analysis (DEA) (Charnes et al., 1978) is a widely used optimization-based technique that measures the relative performance of decision making units that are characterized by a multiple objectives and/or multiple inputs structure. Data envelopment analysis is a technique used to assess the comparative efficiency of homogenous operating units such as schools, hospitals, utility companies, sales outlets, prisons, and military operations. More recently, it has been applied to banks (Haslem, Scheraga, & Bedingfield, 1999) and mutual funds (Haslem & Scheraga, 2003; Galagedera & Silvapulle, 2002; McMullen & Strong, 1998; Murthi, Choi, & Desai, 1997). It is a powerful technique for measuring performance because of its objectivity and ability to handle multiple inputs and outputs that can be measured in different units. The DEA approach does not require specification of any functional relationship between inputs and outputs, or a priori specification of weights of inputs and outputs. DEA provides gross efficiency scores based on the effect of controllable and uncontrollable factors.

The DEA methodology measures the performance efficiency of organization units called Decision-Making Units (DMUs). This technique aims to measure how efficiently a DMU uses the resources available to generate a set of outputs. The performance of DMUs is assessed in DEA using the concept of efficiency or productivity defined as a ratio of total outputs to total inputs. Efficiencies estimated using DEA are relative, that is, relative to the best performing DMU or DMUs (if multiple DMUs are the most efficient). The most efficient DMU is assigned an efficiency score of unity or 100 percent, and the performance of other DMUs vary between 0 and 100 percent relative to the best performance.

Neuro-Fuzzy Inference System Model:

Fuzzy logic starts with the concept of fuzzy sets. Fuzzy sets describe vague concepts. A fuzzy set admits the possibility of partial membership in it. The degree to which an object belongs to a fuzzy set is denoted by a membership function between 0 and 1. A membership function is a curve that describes how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. Fuzzy logic is a convenient way to map an input space to an output space through the primary mechanism of IF-THEN statements called rules.

The input space for the mapping is input parameters and the output space is the decision variables. For instance, the decision maker is advised to accept or reject a proposition or point out the extent of risk involved. Typically, a fuzzy inference system interprets the values of an input vector and, based on some set of rules, assigns values to the output. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned.

Neural fuzzy systems aim at providing fuzzy systems with the kind of automatic tuning methods typical of neural networks but without altering their functionality. In neural fuzzy systems, neural networks are used in augmenting numerical processing of fuzzy sets that is utilized as fuzzy rules. Thus, the fuzzy rule-based modeling process devises a logical approach to imitate the process of human decision making using uncertain information. Neural networks calibrate the model structure to get the optimal model. Neurofuzzy computing optimizes the premise and consequent parameters of the fuzzy inference system using available data. Figure 1 provides a layout of a Neuro-Fuzzy System.

As illustrated in Figure 1, ANFIS takes a fuzzy inference system and tunes it with a backpropagation algorithm based on some collection of input-output data. This allows the fuzzy system to learn. A network structure facilitates the computation of the gradient vector for parameters in a fuzzy inference system. These techniques provide a method for the fuzzy modeling procedure to learn information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given the input/output data. The learning method works similar to that of the neural networks.

A fuzzy inference system works in five steps: fuzzify inputs, apply fuzzy indicator, apply implication method, aggregate all outputs, and defuzzify the output. The first step is to take the inputs and determine the degree to which they belong to each of the appropriate fuzzy sets via membership functions. The input values are always crisp numerical independent variables that measure the operation of the fuzzy system. These crisp variables are then fuzzified via a membership function that computes the measure of belief in the fuzzy indicator. Once all of the inputs have been fuzzified, the inference system combines the fuzzy indicators in ways described by the fuzzy rules in the system. Each fuzzy rule then produces an output value that indicates the measure of certainty in the inferred value. For all rules that produce outputs that relate to the same measure, the centroid of the clipped, fuzzy-membership functions is determined and used to defuzzify the output. This process occurs for all outputs generated by the system

Illustrating DEANFIS model of the Decision Support System for Loan Evaluation

To screen consumer loan applications, loan officers use different methods besides intuitive judgment and experience. Using mathematical techniques, many credit-scoring models have been developed to assist the loan officer in differentiating good loans from bad. Besides these traditional statistical models, many financial institutions use artificial intelligence methods, such as expert systems, artificial neural systems, and fuzzy logic. It is only recently that the finance

community has started applying data envelopment analysis, a relatively new technique. This study proposes to assess the creditworthiness of a new loan applicant using a decision support system that applies a combination of two diverse analytical techniques: data envelopment analysis and neuro-fuzzy models. I propose to use the Data Envelopment Analysis methodology to assess the creditworthiness of an existing set of loans whose outcome (accepted and turned good, accepted and turned bad, and reject) is known. The DEA model benchmarks the given set of loans, and assigns an efficiency score. Further, using the relative efficiency score (generated by the DEA model) and other loan characteristics the neuro-fuzzy model can create rules that identify the major characteristics of loans that are not likely to default.

Each of the loans is a homogenous unit, and we can apply the DEA methodology to asses comparative performance of these loans. The DEA model is a part of a decision support system that uses a number of variables to determine how good a loan is. A loan application includes information such as the applicant's age, housing, address time, total income, number of credit cards, number of dependents, job time, other loan obligations, total debt, monthly rent/mortgage payments, number of inquiries for an applicant, and credit rating. The study proposes to create a DEA model that evaluates the relative efficiency of a set of loans that credit unions have already administered, and allocates a score on the scale of 1 to 100. Further, the next step in the design of the decision support system is to develop a neuro-fuzzy model of the quality of loans analyzed by the DEA model. Artificial neural systems and fuzzy logic are artificial intelligence techniques that function on the same lines as human intelligence. Fuzzy logic is a rule-based development in artificial intelligence that tolerates imprecision and uses it to solve problems. Fuzzy sets and fuzzy logic systems work the same way the brain deals with inexact information. On the other hand, neural networks are modeled after the physical architecture of the brain. Neural networks are specialized hardware or software that emulates the processing patterns of the biological brain. Fuzzy logic and neural networks are complementary technologies in the design of intelligent systems. Each method has merits and demerits. Artificial neural systems suffer from their inability to explain the steps used to make decisions and incorporate rules in their architecture. Neural fuzzy systems address some of the shortcomings of artificial neural network tools. Fuzzy logic techniques often deal with issues such as reasoning on a higher level than neural networks. However, since fuzzy systems do not have much learning capability, it is difficult for a human operator to tune the fuzzy rules and membership functions from the training data set. Thus, to reap the benefits of both fuzzy systems and neural networks, a promising approach is to merge fuzzy logic and neural networks into an integrating system. Neural fuzzy systems represent one of the ways in which fuzzy systems and neural networks can be merged.1 Using a pooled data set of twelve credit unions, this study will design and develop a decision support system (using the combination of DEA and neuro-fuzzy models) to differentiate between "good" loans and potential "bad" loans. Further, the decision support system will also explain why a loan is a bad loan.

Data Specifications:

The data for this study is a pooled data set of loans made by nine different credit unions with a total of 790 2 . The applicants can be categorized into three major groups: applicants who were accepted, and were good credits (Group 1); applicants who were accepted, but were not good credits (Group 2); and applicants who applied for a loan, but were rejected (Group 3). Further, the data set also includes information such as the applicant's age, housing, address time, total income, number of credit cards, number of dependents, job time, co-maker on other loans, total debt, monthly rent/mortgage payments, number of inquiries for an applicant, and credit rating of each applicant.4 Credit unions in the data set assign loan applicants into four credit groups—excellent (1), good (2), marginal (3), and poor (4). The credit rating is determined on the basis of the number of inquiries. The higher the number of inquiries on an applicant, the lower will be the credit rating. The calculation of credit ratings is consistent across all the credit unions. Thus, based on the information supplied by an applicant, we can calculate the applicant's total payments, total income, and total debt.

The study will consider the following variables:

- Total Debt: Total debt of the applicant at the time of application.
- Number of Loans: Total number of loans outstanding in the applicant's name.
- Payments: Total monthly payments
- Dependents: No of dependents of the applicants.
- Total Income: Total monthly income from all sources.
- Job time: Time at current employer.

To evaluate the effectiveness of neuro-fuzzy systems to differentiate between good and bad credit, we use applicant's total income, total payments, total debt, and DEA efficiency score as the input variables. Further, we calculate the ratio of total payments to total income, and ratio of the applicant's total debt to total income.3 If the ratio of total payment to the total income of the applicant is high, the risk of loss due to default by the borrower is high. Similarly, the higher the ratio of total debt to the total income of the applicant, the higher will be the applicant's credit risk. On the other hand, low ratios of total payment to total income and total debt to total income are indicative of a good credit applicant. Besides

[1] The fusion of fuzzy logic and neural networks can be realized in three directions, resulting in systems with different characteristics: Neural Fuzzy Systems, Fuzzy Neural Networks, and Fuzzy-neural Hybrid Systems.

[2] The data set includes loans made by the following nine credit unions: 1. Jefferson County Teachers Credit Union; 2. Jefferson County Employees Credit Union; 3. Family Security Credit Union; 5. Etowah Steel Workers Credit Union; 6. Washington Hill Federal Credit Union; 7. Riverdale Credit Union; 8. Delchamp Credit Union; 9. Lister Hill Credit Union.

[3] Total payments include payments for rent, automobile loan, debt, and other payments. Total income includes income from all sources.

these two ratios, the DEA efficiency score also reflects the creditworthiness of a loan applicant. Therefore, we use three input variables: the ratio of the total payment to the total income (ratio 1), the ratio of debt to the total income (ratio 2), and the DEA efficiency score (calculated credit rating) of the applicant as the factors that can discriminate between a good and a bad loan.4

Data Envelopment Model Specifications for Loan Evaluation

In this study, the analysis of a loan emphasizes inputs and outputs. Therefore, I select the multiplier model for my analysis. In addition, factors such as total debt, number of loans, total payments, number of dependants, total income, and time spent in employment are not very flexible inputs that cannot be immediately controlled. Therefore, output-based formulation is recommended for my study. Furthermore, the quality of the loans does not depend on the scale of operations, thus variable returns to scale is safe assumption. Also, the structure of the DEA model (in multiplier form) uses an equation and separate calculation for every input and output. Therefore, all the input and output variables can be used simultaneously and measured in their own units.

Neuro-Fuzzy Model for the Decision Support System

For the neuro-fuzzy model, I use three input variables: the ratio of the total payment to the total income (ratio 1), the ratio of debt to the total income (ratio 2), and the DEA efficiency score (calculated credit rating) of the applicant as the factors that can discriminate between a good and a bad loan. As mentioned above, the neuro-fuzzy model works in two stages: training and testing. To adequately train the network, the training sample should be a good representative of the population under study. Thus, the training data should cover the entire expected input data space. Further, the network should not be trained completely with input vectors of one class, and then switched to another class; the network will forget the original training. Thus, in accordance with these guidelines, the network should be trained with a sample of 500 observations. The training set is an unbiased sample with data points from all the three classes. Further, to ensure that the training data covered the entire input space (i.e. learn different characteristics of the applications accepted and rejected), observations were selected from all the credit unions. This would prevent the network from learning the characteristics of only one credit union, which can be misleading. Moreover, to ensure that the network is not trained with vectors from one class or one single credit union, the observations should be intermingled randomly. Finally, as there are no preferable membership functions, I will create an initial set of membership functions using grid partition method. The built-in function genfis 1 of the fuzzy logic toolbox of the MATLAB software will be used to create the initial membership function matrix.

CONTRIBUTION OF THE STUDY

This study proposes the modeling and development of a decision support system that uses a combination of data envelopment analysis and neuro-fuzzy systems. Thus, the decision support system derives benefit from both methodologies to recommend a decision. As illustrated in the literature review section, many studies illustrate the use of fuzzy systems and DEA. However, very few studies illustrate the fusion of DEA and ANFIS models. In addition, all the studies illustrate fuzzy DEA models (using fuzzy logic and DEA model). This study proposes the design and modeling of DEANFIS model (using DEA and ANFIS model) to develop a multidimensional decision support system that can benefit from both the techniques.

DEA does not require the manager to attach prescribed weights to each input and output. Moreover, DEA modeling does not require prescription of the functional forms that are needed in statistical regression approaches. DEA uses techniques such as mathematical programming that can handle a large number of variables and constraints. As DEA does not impose a limit on the number of input and output variables to be used in calculating the desired evaluation measures, it's easier for managers to deal with complex problems and other considerations they are likely to confront. DEA identifies good units in a given set of DMUs and provides a measure of inefficiency for all others. The DMUs having the most desirable characteristics are rated a score of one (100% efficient), while the DMUs that are inefficient score between zero and one. DEA methodology can identify a bad DMU by comparing its characteristics with a given set of benchmark DMUs having good DMU characteristics.

Similarly, neuro-fuzzy models do not require restrictive assumptions of the statistical model. Fuzzy logic provides a means of combining symbolic and numeric computations in inference processing. The linkage between neural networks and symbolic reasoning can be established through the membership function of fuzzy logic. The membership function measures the degree of possibility of a concept as related to a numeric quantity. A neural network can be used to synthesize a membership function by training it with instances of the relation. Neuro-Fuzzy systems provide flexibility to the decision-maker to incorporate their own rules in the DEA model to assess DMUs.

FIGURES, TABLES, & REFERENCES

Available upon request from the author.

[4] According to Fabozzi, in a mortgage loan, banks and financial institutions consider the ratio of total payments to the total income and the ratio of debt to the total income of the applicant. Since consumer loans are somewhat similar in nature, we use these ratios to discriminate between good and bad credit risk. See Fabozzi (1996).