A Neuro-fuzzy System with Semi-supervised Learning for Bad Debt Recovery in the Healthcare Industry

Donghui Shi  
Department of Computer Engineering  
School of Electronics and Information Engineering  
Anhui Jianzhu University  
Hefei, China 230601  
sdonghui@gmail.com

Jozef Zurada  
Department of Computer Information Systems  
College of Business  
University of Louisville  
Louisville, KY 40292  
jozef.zurada@louisville.edu

Jian Guan  
Department of Computer Information Systems  
College of Business  
University of Louisville  
Louisville, KY 40292  
jeff.guan@louisville.edu

Abstract

The cost of healthcare in the U.S. has been rising with an alarming rate during the last two decades. One of the causes for the rising cost is medical bad debt. There is surprisingly very little research that explores the performance of computational intelligence and soft computing methods in improving bad debt recovery in the healthcare industry. This study examines the performance of an adaptive neuro-fuzzy inference system (ANFIS) with semi-supervised learning (SSL) in classifying bad debt in the healthcare context, as better debt classification leads to greater recovery. Computer simulation shows that ANFIS with SSL is a viable method. Our model generated better classification accuracy than ANFIS alone did and our accuracy is better than those in the previous studies. Insightful interpretation of the results is provided through data clustering and analysis of control surfaces. The latter depicts nonlinear interaction between various factors contributing to bad debt. Additional analysis is provided through receiver operating characteristic (ROC) charts to interpret the classification accuracy rates at a continuum of cut-off points from within the range [0,1]. These results and their analysis show the potential of ANFIS with SSL models in classifying unknown cases, which are a potential source of revenue recovery.

1. Introduction

In 2011 the United States spent on healthcare 17.2% of its gross domestic product, $8,608 per capita, significantly more than any other industrialized nation. Moreover, the cost of healthcare has steadily been rising during the last two decades. One of the factors that contribute to the high costs is medical bad debt. This includes unpaid patient bills for medical treatment, outstanding medical testing costs, and collection agency fees [4],[7],[10],[15]. According to Galloro [4] Nashville-based Hospital Corporation of America’s provision for bad debt rose to an astounding 10.3% of its net revenue for just one quarter, compared to an already high 8.3% of revenue in the same quarter the previous year. Even when unpaid bills are eventually paid, hospitals typically end up paying 30% to 50% of recovered bad-debt revenue to outside collection agencies [15]. More recently the rising uninsured, underinsured, and patients enrolled in high-deductible plans are adding to the budget pressures [1]. The bad debt issue in healthcare is not only affecting the bottom line, but it also has an impact on a healthcare organization's ability to provide care [10]. According to an analysis from Citi projects bad debt could reach $200 billion by 2019 [7]. Recovering bad debt has become a serious matter and has led to hospitals suing patients in several states in the U.S. According to information provided by collection attorneys, consumer advocates, and court records, some hospitals use extreme measures to collect bad debt and even seek arrest of patients who miss court hearings related to their healthcare debts [8].

There is a great deal of research on scoring, managing, and recovering distressed debts from a loan, a credit line, or an accounts receivable. For example, Murgia and Sbrilli [9] compare the performance of a neural network, integer programming, hidden Markov model, logistic regression, regression tree, and Bayesian classifier in distressed debts recovery. The authors report that the neural network scored the recovery rate of each distressed debt better than the other models. However, despite an increasingly obvious and urgent need for predictive and classification models of bad debt in the healthcare industry, academic research on this very important topic appears to be surprisingly scarce. A thorough literature search on this topic brought only a few references summarized below.

An early study by Zollinger [16] examines a sample of 985 patients from 28 Indiana hospitals using a regression model and identified several institutional variables, such as total hospital charge and the total hospital revenue, and patient variables, such as marital
status, gender, diagnoses, insurance status, employment status, and discharge status, are significant factors in recovering unpaid hospital bills. Buczeko [2] analyzes data on charges assigned to bad debt for 82 short-stay hospitals in Washington. The author confirms that unpaid care has become a serious problem in hospital finance because of the increasing number of uninsured patients and declining hospital revenues. Veletsos [15] describes a more comprehensive study on using predictive modeling software such as IBM Intelligent Miner and DB2 for bad-debt recovery. The study includes approximately 2400 patients of the Florida Hospital in Orlando. The final model is based on a variety of data variables, including credit factors, demographic information, and previous organizational payment patterns. The model yields approximately $200,000 in savings. Pesce [11] argues that hospitals should invest in modern information technology to reduce bad-debts, which along with other factors such as billing errors, insurance underpayments, and inability to collect accurate patient and payer information throughout delivery of care, account for 13% of a hospitals' lost revenue each year. More recently a report describes the use of predictive software in IBM SPSS to improve the bad debt collection effort and boost revenue. Though no details (such as models and methods used) are provided, the report states that "one hospital saw a 30% reduction in bad-debt write-offs, a 12 percent increase in self-pay collection rates, and $25,000 per month reduction in agency fees" [5].

Zurada and Lonial [17], [18] compare the performance of several computational intelligence tools in recovering bad debt. Their results show that the logistic regression, neural network, and the ensemble models produce the best overall classification accuracy, and the decision tree is the best in classifying cases for which the debt has been recovered. The models are also used to score the "unknown" cases – those that were not pursued by a company. The neural network model classifies more "unknown" cases into "good" (recoverable) cases than any other models tested in the studies. This may potentially provide an additional source of recoverable income. Finally, the paper by Shi et al. [12] examines the effectiveness of a neuro-fuzzy method (ANFIS) under numerous scenarios in classifying bad debt. In the classification accuracy rates and scoring of "unknown" cases ANFIS outperforms the methods used in the studies by Zurada and Lonial [17], [18].

Predicting whether a particular customer is likely to repay a healthcare debt is an inherently complex and unstructured process. What makes this process especially difficult in the healthcare context, especially in emergency rooms and acute care facilities, is their inability to obtain detailed financial information concerning the patients. Unlike a financial institution which would collect financial, social, and personal information about a customer and carefully evaluate whether to extend him/her a loan, healthcare institutions must often admit a patient and perform the necessary medical procedures on credit knowing very little about the particular patient. This lack of information makes it difficult to predict whether a patient-debtor will pay his/her bill or not. Thus, due to moral, legal and practical constraints, healthcare providers in the U.S. often become unwilling creditors to a multitude of borrowers. According to Pesce [11] a healthcare institution is handicapped by having only a small number of independent attributes of the patient-debtor to evaluate. In addition, some debt defaults may be attributed to unforeseen events (i.e. divorce, death, accident, disability, loss of employment) or be governed by factors that may be difficult or impossible to detect in the attributes of the consumer (i.e. stability of marriage, general health, job stability). Given all of the above-mentioned reasons it is not surprising that the healthcare institution that provided data for this study recovered bad debts from only 7.3% of the non-paying patients.

This paper examines the effectiveness of a neuro-fuzzy method (ANFIS) with semi-supervised learning (SSL) in classifying bad debt. It also compares the classification accuracy rates to two previous studies [12], [18]. The target/dependent variable in a fairly large data set provided by a healthcare institution represents the following three classes: 1: "good" customers (those who repaid the debt or made partial payments to repay the debt); 2: "bad" customers (those who defaulted or refused to repay the debt); and 3: "unknown" customers (those who were not pursued). Due to the low recovery rate, the number of "good" customers is vastly underrepresented in the data set. To build the initial ANFIS model we only used cases representing "good" and an equal number of randomly selected "bad" customers. The initial model was then used in semi-supervised learning to classify the unknown cases. The labeling of "unknown" cases into "good" or "bad" provides a potential source of additional revenue to the company.

We ran computer simulation and used different membership functions for ANFIS. The interpretation of the control surfaces generated by ANFIS gave unique and preliminary insights into the interactions between the input variables and the probability of default/recovery. The paper is organized as follows. Section 2 describes the data sample and section 3 presents the basic characteristics of ANFIS with SSL used in the study. Section 4 discusses the experiments and simulation results as well as compares them to two previous studies. Finally, section 5 concludes the paper and provides recommendations for future work.
2. The data sample

The healthcare company, whose data were used in this study, relied on only four simple input factors to determine whether a bad debt was recoverable: (1) Patient Age (PA), (2) Patient Gender (PG), (3) Injury Diagnosis Code (IDC), and (4) Dollar Amount of the Claim (DAC). In all likelihood, the four factors constituted all of the information about the patient-debtor that was available to the healthcare company. Furthermore, aside from the amount owed, the information appears to be only tangentially related to the probability that a particular bad debt could be recovered. The dataset contains 6117 cases with an outstanding balance of $2,381,453. The dependent variable, Status, comprises of 449 "good" cases (group 1), 2833 "bad" cases (group 2), and 2835 "unknown" cases (group 3).

To learn more about the distribution of the variables within the data set and to find out whether any transformation of the variables was needed, we calculated the descriptive statistics. The results are summarized in Table 1. Out of 24 IDC groups numbered 1-24, Table 1 presents the five most typical IDC groups only – those that occur the most frequently. These five IDC groups account for about 76% of all IDC codes. The remaining 19 IDC groups are not shown. The table shows that for the DAC variable the average dollar amount of the recovered cases (group 1), not recovered (group 2), and not pursued (group 3) are $1,052, $417, and $256, respectively. Furthermore, the table shows that the total amounts for the DAC variable for each of the 3 groups are $472,461, $1,182,142, and $727,850, respectively. Thus it appears that the company used common sense and some procedure that allowed it to target the patients with larger debt amounts and ignore those with smaller debt amounts. Furthermore, the skewness coefficient $S_k=19.1$ shows that the distribution of the DAC variable is very positively skewed for group 3, which suggests that small dues were simply not pursued. For cases belonging to group 3 ("unknown" cases) debt collection was not pursued by the subject company, but these unknown cases may represent a potential source of debt recovery. It is possible that the low debt recovery rate might have been caused by the healthcare company using primarily the amount owed to determine which bad debts to target. The purpose of our data mining techniques is to use the seemingly unrelated
factors such as the patient’s gender, age and type of injury to determine the likelihood that a particular patient-debtor will pay his/her overdue bill. To build the initial ANFIS model we only used cases representing "good" and an equal number of randomly selected "bad" cases. The initial model was then used in semi-supervised learning to classify the unknown cases. The labeling of "unknown" cases into "good" or "bad" could provide additional revenue to the company. To improve the distribution of the DAC variable and obtain better prediction results, we used $\log_{10}(\text{DAC})$ instead of DAC.

3. An adaptive neuro-fuzzy inference system with semi-supervised learning for bad debt recovery

In this section we describe the basic principles of semi-supervised learning and ANFIS. Semi-supervised learning leverages both labeled and unlabeled data examples. It aims to combine unsupervised learning, which utilizes no data labels, and supervised learning, which utilizes data for which labels are present. The learning algorithm generally uses a smoothness assumption, which states that if two examples are relatively close in feature space, then their corresponding class outputs should be close in class space. In a semi-supervised learning a classifier is iteratively built on its own predictions. First, a classifier is built on the labeled data and used to classify unlabeled data. Typically the most confidently predicted examples are then iteratively inserted into the training set and a new classifier is generated [3]. A classifier could be implemented using $k$-nearest neighbor method, support vector machines, neural networks, logistic regression, or a neuro-fuzzy system. In this paper we employ an ANFIS-based classifier.

A general self-training algorithm for the imbalanced data follows.

Input: $L_I$ – the original labeled data set; $U$ - unlabeled data set
Output: $L_F$ – final labeled data set with all cases from $U$ classified as good or bad
Repeat:
1. Merge all good cases from $L_I$ and the same number of bad cases randomly selected from $L_I$ to produce a new dataset $D$ with balanced class labels
2. Use $D$ to construct an ANFIS-based predictive model $F$
3. Classify unlabeled data set $U$ with $F$
4. Compute the error
5. Select case $u$ with the minimum error and move the case from $U$ to the $L_F$: $L_F = L_F + u$; $U = U - u$
Until $U$ is null

Originally the labeled data set $L_I$ contained 3282 cases of which 449 were good cases and 2833 were bad cases. Unlabeled data set $U$ contained 2835 cases. After employing ANFIS with SSL, the final data set contained 6117 cases of which 1657 and 4460 were classified as good and bad cases. There were no unlabeled cases as they were classified either as good or bad cases.

The block diagram of ANFIS with SSL is depicted in Figure 1.

![Figure 1. The diagram of ANFIS with SSL](image)

Neural fuzzy inference systems have emerged from the fusion of artificial neural networks and fuzzy inference systems. These systems combine learning/training and optimization abilities of artificial neural networks with human-like reasoning using if-then fuzzy rules offered by fuzzy inference systems. Neuro-fuzzy inference systems have formed a popular framework for modeling real world problems including...
classification. ANFIS is one of the better known neuro-fuzzy inference systems [6]. One of the advantages of ANFIS is its ability to generate fuzzy sets represented by membership functions and fuzzy rules from preexisting input-output data pairs available in the data set. Figure 2 shows the architecture of the ANFIS bad debt classification model in this paper. The model has 4 (m) inputs representing the 4 patient characteristics described in section 2. Each of the inputs has 2 (n) membership functions. The model uses a typical ANFIS architecture with an additional node at the output end representing a discrimination function that classifies the output as either “good” customer (debt repaid) or “bad” customer (debt unpaid) with a user specified threshold value. The model uses a Takagi, Sugeno, and Kang (TSK) type fuzzy inference system and has two sets of trainable parameters: the antecedent (premise) membership function parameters and the consequent (polynomial) parameters [13], [14]. A typical TSK rule has the following structure:

\[ f = r + p_1 X_1 + p_2 X_2 + \ldots + p_m X_m \]

where \( A_{ij} \) is the \( j \)th linguistic term (such as high, low) of the \( i \)th input variable \( X_i \), \( m \) is the number of inputs, \( f \) is the estimated output, and finally \( r \) and \( p_i \) are the consequent parameters to be determined in the training process. The architecture in Figure 2 is described as follows:

**Figure 2. Architecture of the ANFIS bad debt recovery model**

**Layer 1:** This layer contains the membership functions with adaptive parameters or premise parameters. The number of nodes (\( N=8 \)) in the first layer is the product of the input size (\( m=4 \)) and the number \( n=2 \) of the membership functions for each input variable, or \( N=m \times n \). The output of each node is defined as

\[ O_{ij} = \mu_{ij}(X_i) \]

where \( \mu_{ij} \) is the \( j \)th membership Gaussian function (four other functions have been used in this study) for the input \( X_i \) and is given as follows:

\[ \mu(X) = \exp \left\{ - \left[ \frac{x-c}{a} \right]^2 \right\} \]

where \( a \), \( b \), and \( c \) are the premise parameters.

**Layer 2:** This layer calculates the firing strength of each rule and the output in this layer represents these firing strengths. The output is the product of all of its inputs as follows:

\[ O_k = W_k = \mu_{11}(X_1) \mu_{21}(X_2) \cdots \mu_{m1}(X_m) \]

for \( k=1,R \) and \( R \) is the number of rules.

**Layer 3:** This layer normalizes the weighing factor of each of the input nodes \( k \) as follows:

\[ O_k = \overline{W}_k = \frac{W_k}{W_1 + W_2 + \cdots + W_R} \]

**Layer 4:** the output of this layer represents a weighted value of the first order fuzzy if-then rule as follows:

\[ O_k = \overline{W}_k f_k \]

where \( f_k \) is the output of the \( k \)th fuzzy rule as follows:

\[ f = r + p_1 X_1 + p_2 X_2 + \ldots + p_m X_m \]

where \( p_i \) and \( r \) are called the consequent parameters and \( j = 1, \ldots, m, k = 1, R \).

**Layer 5:** Finally this single node layer computes the overall output \( (F) \) of the ANFIS model as the sum of all the weighted outputs of the previous layer as:

\[ O = F = \sum_{k=1}^{N} \overline{W}_k f_k \]
where $f_k$ represents the output of the $k^{th}$ TSK-type rules as defined in layer 4.

The last module is a discriminant function $f(F)$ which receives $F$ as input and maps it to output $C$ which is one of two values, “good” customer or “bad” customer. The parameters, both the premise parameters and consequent parameters, are learned/optimized in the training process. Two parameter optimization methods are used in training. The first method is backpropagation and the second method is a hybrid method that uses a mixture of backpropagation and least squares.

4. The experiments and simulation results

Computer simulation was conducted using MatLab Fuzzy Logic Toolbox. 50% of the records in the data set were randomly allocated to building the models, whereas 25% of the records were allocated to the models’ validation and testing, respectively. Since classification accuracy rates may vary significantly for different partitions/splits of the data set, this process was repeated 50 times and the reported classification rates on the test sets were averaged over the 50 runs to eliminate the classification bias resulting from random splits of the data set and to increase the reliability and generalizability of the results. We used the backpropagation method for training the fuzzy inference system (FIS) membership function parameters and GENFIS1 function to generate the initial FIS. We used two membership functions per input variable and five types of different membership functions. These are two Gaussian membership functions (gauss2mf and gaussmf), generalized bell-shaped membership function (gbellmf), the difference between two sigmoid membership functions (dsigmf), and triangular membership function (trimf). To compare the models’ performances, we used the overall correct classification accuracy rates as well as the rates for good and bad cases.

We also utilized the ROC charts, which depict the global performances of the models within the [0,1] range of cutoffs. Low and high probability cutoffs tend to be in the upper right and lower left areas, respectively, of the ROC curves. To interpret the results we also used 3-dimensional control surfaces generated by ANFIS for the unclustered data and data clustered by most frequently occurring IDCs. We also compared the classification accuracy rates to those presented in scenario 3 of the Shi et al. study [12] and Zurada and Lonial study [18]. In both of those studies and in this study the same data set was used.

Table 2 depicts the classification accuracy rates from this study. It seems that a choice of the membership function affects the rates. The overall rates are between 74.8% and 76.7% and the best rates came from the model using the generalized bell-shaped membership function (gauss2mf). The good rates are within the [80.9%, 84.0%] range and the bad rates are within [66.9%, 70.9%] range. ANFIS with SSL generated better results than those obtained from ANFIS alone as reported in [12] (Table 3) as well as those results from the five models (decision tree, neural network, logistic regression, memory-based reasoning, and ensemble) reported in [18] (Table 4). The improvement in the overall rates could be mainly attributed to the increment in the rates for good cases.

| Table 2. The classification accuracy rates in [%] for cutoff=0.5 – ANFIS with SSL |
|----------------------------------|------------------|------------------|------------------|------------------|------------------|
|                                  | gauss2mf         | gbellmf          | dsigmf           | gaussmf          | trimf            |
| Overall                          | 76.7             | 76.3             | 74.8             | 75.9             | 75.5             |
| Good                             | 84.0             | 83.1             | 82.6             | 80.9             | 81.5             |
| Bad                              | 69.3             | 69.6             | 66.9             | 70.9             | 69.3             |

| Table 3. The classification accuracy rates in [%] for Scenario 3 for cutoff=0.5 – ANFIS alone [12] |
|----------------------------------|------------------|------------------|------------------|------------------|------------------|
|                                  | gauss2mf         | gbellmf          | dsigmf           | gaussmf          | trimf            |
| Overall                          | 61.6             | 63.7             | 62.8             | 63.1             | 62.5             |
| Good                             | 61.4             | 63.5             | 61.9             | 63.9             | 63.5             |
| Bad                              | 61.9             | 64.2             | 64.0             | 62.7             | 62.0             |

| Table 4. The classification accuracy rates in [%] for cutoff=0.5 – [18] |
|---------------------------|-----------------|-----------------|-----------------|-----------------|
| Decision Tree             | Neural Network  | Logistic Regression | Memory-based Reasoning | Ensemble Model |
| Overall                   | 67.9            | 72.3            | 75.0            | 61.2            | 73.7           |
| Good                      | 75.0            | 67.9            | 71.4            | 72.3            | 71.4           |
| Bad                       | 60.7            | 76.8            | 78.6            | 50.0            | 75.9           |

One can compare the above observations by examining the areas under the ROC charts in Figure 3. While there is little difference among the overall performances of the five models (five different membership functions) in the current study, all five models significantly outperform the model from [12].
Table 5. Summary of the descriptive statistics after semi-supervised learning

<table>
<thead>
<tr>
<th>Status</th>
<th>Patient Gender (PG)</th>
<th>Patient Age (PA)</th>
<th>Most frequently occurring Injury Diagnosis Code (IDC). Code Range</th>
<th>Dollar Amount of the Claim (DAC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (groups 1,2)</td>
<td>Female (N=2884, 47.1%)</td>
<td>Male (N=3233, 52.9%)</td>
<td>Mean=30 St. Dev=21 Min=0 Max=100 Sk=0.7 &quot;800&quot;.&quot;829&quot; 16 &quot;840&quot;.&quot;848&quot; 18 &quot;870&quot;.&quot;897&quot; 14 &quot;920&quot;.&quot;924&quot; 9 &quot;958&quot;.&quot;959&quot; 19</td>
<td>Mean=$389 St. Dev=$1,477 Min=$1 Max=$40,508 Sum=$2,381,453 Sk=12.8</td>
</tr>
<tr>
<td>1 (Good) N=1657</td>
<td>Female (N=969, 58.5%)</td>
<td>Male (N=688, 41.5%)</td>
<td>Mean=23 St. Dev=18.8 Min=0 Max=100 Sk=1.1 &quot;800&quot;.&quot;829&quot; 19 &quot;840&quot;.&quot;848&quot; 22 &quot;870&quot;.&quot;897&quot; 15 &quot;920&quot;.&quot;924&quot; 9 &quot;958&quot;.&quot;959&quot; 6</td>
<td>Mean=$616 St. Dev=$2,281 Min=$1 Max=$40,508 Sum=$1,020,834 Sk=10.4</td>
</tr>
<tr>
<td>2 (Bad) N=4460</td>
<td>Female (N=1915, 42.9%)</td>
<td>Male (N=2545, 57.1%)</td>
<td>Mean=32.4 St. Dev=21.6 Min=0 Max=100 Sk=0.6 &quot;800&quot;.&quot;829&quot; 14 &quot;840&quot;.&quot;848&quot; 17 &quot;870&quot;.&quot;897&quot; 13 &quot;920&quot;.&quot;924&quot; 9 &quot;958&quot;.&quot;959&quot; 24</td>
<td>Mean=$305 St. Dev=$1,016 Min=$1 Max=$19,568 Sum=$1,360,619 Sk=9.5</td>
</tr>
</tbody>
</table>

Table 5 represents the descriptive statistics for the resulting cases after labeling the unknown cases with ANFIS using SSL. The format of the table is very similar to that in Table 1. There are no unknown cases as they have already been classified as good cases or bad cases. Comparing Table 1 with Table 5 yields some interesting findings. Females are more likely to repay the medical bad debt than males. It also appears that younger patients are less likely to default on their bad debt. Also, it is much more likely to recover the bad debt for medical procedures represented by the following IDCs: "800"."829" (Fractures) and "870"."897" (Open Wounds) and less likely for "920"."924" (Contusion w/Intact Skin) and "958"."959" (Complications & Unspecified Injuries). There does not appear to be an obvious explanation for this last set of results based on IDCs. However, further investigation into the diseases, disorders, and/or symptoms by a healthcare organization may shed light on these interesting findings. The 1208 unknown cases which were classified as good cases can potentially bring ($1,020,834-$472,461)=$548,373 in additional revenue (Tables 1 and 5). It also seems that the models learn to pursue larger debts.
Figure 4. Six control surfaces

Figure 4 depicts six 3-dimensional surfaces generated by ANFIS with SSL. They offer more insight into nonlinear and complex interactions between the variables and the probability of bad debt default/recovery. The surfaces have been plotted for the gbell membership function as this function generated stable classification rates comparable to the rates produced by the gauss2 membership function. For example, the chart log(DAC) vs. (IDC) indicates that patients who owe smaller amounts and have injuries represented by higher IDC codes are more likely to default. Apparently the healthcare company has not been interested in pursuing small debts. From the log(DAC) vs. PA plot one can conclude that it is much more difficult to recover bad debt from older patients. The remaining charts show that males are more likely to default than females. One can see that all six charts generally confirm the observations described earlier. The control surfaces in Figures 5 and 6 represent the relationships between log(DAC), PA, and PG for the data clustered by the most frequently occurring IDCs. Out of the five IDC groups (Table 1 and Table 5), we selected two example IDC groups, i.e., Fractures represented by IDCs "800"-"829" and Complications & Unspecified Injuries represented by IDCs "858"-"859". As discussed earlier, it appears that it is more probable to recover the bad debt for the former IDC group than the latter. The surface log(DAC) vs. PA in Figure 5 indicates the highest probability of default (the peak) occurs for the middle age patients with relatively low values of DACs. And the chart log(DAC) vs. PG appears to be somewhat counterintuitive as it seems that patients with smaller DACs are more likely to default. The surface PA vs. PG confirms the previous findings.
that it is more probable that males and older patients are not going to repay their debt. The three control surfaces in Figure 6 can be interpreted in a similar manner.

Figure 5. Three control surfaces for IDC=4 ("800"-"829": Fractures)

Figure 6. Three control surfaces for IDC=18 ("958"-"959": Complications & Unspecified Injuries)

5. Conclusion

The paper explores the effectiveness of ANFIS with semi-supervised learning (SSL) in recovering bad debt in the healthcare context and compares its performance to two previous studies. The data analysis and evaluation of the performance of the models is based on a fairly large unbalanced data sample provided by a healthcare company, in which cases with recovered bad debts are underrepresented. This research was motivated by an urgent and recognized need to better understand the effectiveness of computational intelligence and soft computing methods in bad debt recovery in the healthcare industry and relatively low level of academic interest in this field. This paper describes a study that explores the effectiveness of ANFIS with SSL in classifying bad debts. The results indicate that SSL could potentially lead to a higher bad debt recovery rate through classification of unknown cases. Five different models were designed and tested. The results and the approach in this study could potentially help health-care organizations target specific
group of customers to improve their return on debt recovery efforts. The example control surfaces for all data and the data clustered by IDCs reveal interesting relationships between the probability of recovery/default and the two other variables. Though there does not appear to be an obvious interpretation for some interesting results, further investigation into the diseases, disorders, and/or symptoms by a healthcare organization may shed more light on these interesting findings.

Finally the paper also shows the ability of the models to classify unknown cases, which are a potential source of revenue recovery. This study shows the potential of computational intelligence and soft computing models to classify bad debts using data sets whose features contain very tangential information about the patients. An important fact about the dataset used in the study is the high proportion of unknown cases. These unknown cases often represent huge amounts of possible recoverable revenue. Further research can examine the interpretable rules generated by ANFIS and specific IDCs to find out if patients who suffered from a more serious injury (such as those that may lead to disability) are less likely to pay the debt than patients who were treated for minor injuries. Our results indicate likely existence of clusters in both DACs and IDCs. One possible future research direction is to explore further improvement in classification through clustering of data by DAC and/or IDC.

Acknowledgment: We express gratitude to Professor Subhash Lonial, Ph.D., Department of Marketing, College of Business, University of Louisville, for providing the data set used for computer simulation in this paper.

6. References


