Transparent Decision Support System for Credit Risk Evaluation: An automated credit approval system

Abhinaba Dattachaudhuri
Computer Science and Engineering
National Institute of Technology
Silchar, India
ad.chaudhuri1995@gmail.com

Saroj Biswas
Computer Science and Engineering
National Institute of Technology
Silchar, India
bissarojkum@yahoo.com

Sunita Sarkar
Computer Science and Engineering
Assam University
Silchar, India
sunitasarkar@rediffmail.com

Arpita Nath Boruah
Computer science and Engineering
National Institute of Technology
Silchar, India
arpita.boruah@hotmail.com

***Abstract*—The assessment of financial credit risk is an important and challenging research topic in the financial domain. The accurate assessment and prediction of financial credit risk play an essential role both on economics and society. Many banking sectors are looking into an automated decision support system for credit risk evaluation. Consequently, many automated credit risk evaluation systems have been proposed, however most of them lack in explainability needed to justify their decisions. To remedy this, this paper proposes a decision support system named Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE), which extracts concise and justifiable rules from neural networks for credit risk decision. The proposed TDSSCRE tunes and prunes the rules to assist in making a decision support system transparent. Finally this transparent decision support system with concise rules justifies the decision for why applications are granted/rejected with a significant predictive accuracy. The decision system is validated with 3 credit risk datasets using 10 fold cross validation. From the experimental results it is observed that the proposed decision system can produce accurate and justifiable rules for classification. The decision system is also compared against two existing rule extraction models.**

***Keywords*—** *Decision System, Credit Risk, Neural Network, Machine Learning, Rule Extraction*

I. INTRODUCTION

Historically the process of granting a loan is an important task and highly debated topic in the financial domain. Until recently the decision of granting loan is a manual process by inspecting an applicant's personal information. A credit expert takes into account an applicant's personal information - social status, economic status, character history and other factors to decide the applicant's creditworthiness. Since last decades have brought with them huge advancements in digital communication and storage, there is a trend of archiving all sorts of data relating to any business organizations. Financial institutions are known to collect and compile huge volumes of transactional data every day. This huge volume of data leaves no room for individuals [1] to extract hidden information from the data, however with the advent of machine learning concept, there are some intelligent machine learning algorithms which can process and analyses the data deeply to mine some meaning information from the financial data. Therefore, many banking sectors are looking into an automated and efficient machine learning method for credit risk evaluation. Consequently, many statistical and predictive models have been proposed for credit risk evaluation [2]. These models belong to different categories namely statistical methods, nonparametric methods and Neural Network (NN) models. Purohit et al. [3] have shown the applicability of an integrated model on financial data from Indian banks. The combination model is based on Logistic Regression, Multilayer Perceptron, Radial Bias Neural Networks, SVM, C4.5 Decision Tree. They have shown that the performance of MLP Neural Network is the best in many cases where the data has missing values. Khandani et al. [4] have applied the method of generalized Classification And Regression Trees (CART) for the classification purpose of the financial data. The development of these models is aimed towards achieving the highest predictive accuracy, without any explanation of how the decision is made. However, in many cases, a financial organization needs to justify the reason why a credit request is accepted or declined and hence the predictive model needs to be explainable. Credit-risk evaluation systems should focus more on providing explanations for why applications are rejected [5] instead of merely trying to improve predictive accuracy. By providing justification for their decision financial organizations can improve customer relations. This in turn will draw in more customers, benefiting both the customer and the financial organizations. Taking into account the above issues, this paper proposes a decision support system named Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE) which proposes a transparent rule extraction algorithm using neural network for rule extraction from financial data. The proposed model generates rules using neural networks because neural networks are powerful classifiers and have a universal

978-1-7281-4994-3/20/$31.00 ©2020 IEEE

approximation property which makes them very useful for classifying financial data [6-7]. The generated rules are tuned and pruned to assist in making a decision support system transparent. Finally this transparent decision support system with concise rules justifies the decision for why applications are granted/rejected with a significant predictive accuracy.

II. LITERATURE SURVEY

A lot of work has been done on the issue of automated credit risk analysis. Addo et al. [8] compared the performance of four models for the credit risk prediction. The models used are, Elastic Net, Random Forest Modeling, Gradient Boosters, and Deep Learning Models. They have shown that overall performance is the best for Gradient Booster and Deep Learning Models. Purohit et al. [3] have used data source from Indian Banks to show the applicability of an integrated model on financial data. The combination model comprises of Multilayer Perceptron, Logistic Regression, Radial Bias Neural Networks, SVM, C4.5 Decision Tree. Their results show that MLP neural networks have an edge in classifying data that have inherently missing values. Khandani et al. [4] have applied the method of generalized Classification And Regression Trees (CART) for the classification purpose of the financial data. The CART model produces comprehensible rules from trees but cannot compensate for missing data. Chen et al. [9] have shown a comparative review to commonly used decision models. They conclude that decision systems should focus on some credit scoring scheme that can obtain the degree of credit risk along with a prediction.

There have been many proposed rule extraction algorithms that reveal the information contained in the neural networks. These approaches have three types namely, pedagogical, decompositional and eclectic [10]. The decompositional techniques analyze the weights and activation functions of the layers to compute rules that best resemble the network. Setiono et al. [11] proposed a decompositional algorithm called the Neuro Rule algorithm that extracts classification rules from a single hidden layer of neural networks. The rule generation component of Neuro Rule [11] is called RG which generates rules such that it can cover the maximum number of examples with a minimum number of attributes. Biswas et al. [12] have proposed the RxNCM algorithm which uses misclassified patterns after removing insignificant neurons to generate the rules. The RxNCM [12] algorithm uses both the misclassified as well as the properly classified algorithm to generate the rules. There also exists eclectic approaches that combine decompositional with pedagogical approach. The Rex-CGA method proposed by Hruschka et al. [13] works with multiple layers of hidden layers. Clusters of activation values in the hidden layers are found by using clustering genetic algorithm, with which Rex-CGA generates rules. The DEDEC [14] method proposed by Tickle et al. extracts rules by identifying the minimal set of information required to distinguish individual patterns. Pedagogical techniques have been shown to be less computationally demanding while having simplicity in their implementation and higher accuracy than other approaches [15]. For the reasons of simplicity and for being computationally light, this paper adopted the pedagogical approach for the credit risk assessment system.

III. THE PROPOSED TDSSCRE MODEL

The TDSSCRE decision system generates the rules for explanation and classification. TDSSCRE system extracts rules using neural network. TDSSCRE system follows the following steps to generate the decision rules: *Network Training, Network Pruning, Data Range Computation, Rule Generation, Rule Pruning and Updation, Final Rule Set.* The optimal network architecture is found using the back-propagation algorithm. The TDSSCRE algorithm analyses both improperly classified and correctly classified patterns to compute the data ranges. These data ranges are used to generate the rule set required for classification and explanation.
The overview of the TDSSCRE is given below in Figure 1.

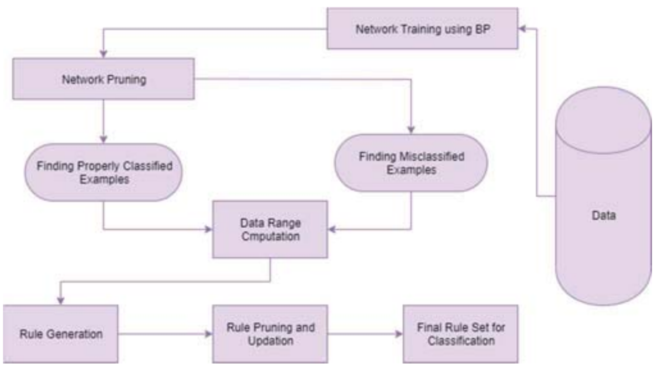

Fig. 1. Overview of TDSSCRE

Network Training:

A back propagation NN with single hidden layer and ***h*** hidden neurons is used for rule extraction. The number of *hidden neurons* is selected based on the mean square error of the network. The network architecture with ***l*** number of input neurons is varied from ***l+1*** to ***2*l*** hidden neurons and the architecture with least mean square error is chosen for further processing

Network Pruning:

The Neural Network is pruned by finding the significance of the input neurons. At each iteration of this step, the error produced by an input when it is removed from the network is calculated by mean square error. The input whose removal causes the network to have the least error is considered insignificant. At the end of each iteration insignificant inputs are removed from the network.

The set of insignificant inputs (say, B) is conditionally included back into the network to see if there is any improvement in the test accuracy or not. The concept of sequential floating search is applied to prune the network.

Data Range Computation:

The data range of each remaining neuron is calculated using both misclassified and properly classified patterns. For each input neuron, find misclassified rate by removing that neuron and finding the misclassified patterns. Find properly classified rate by finding properly classified patterns using only that neuron. The union of the misclassified and properly classified patterns is used to compute the data matrix. The data range is found using the upper and lower limits of the patterns found.

Rule Generation:

The rule generation phase extracts the mandatory attribute ranges from the data ranges obtained to constructs the classification rules. The rules are formed using the upper (U_{ik}) and lower (L_{ik}) limit as shown:
if(attribute1>=L_{ik} and attribute1<=U_{ik}) then class = 'A' else class = 'B'

Rule Pruning and Updation:

These rules are pruned to remove insignificant ones in the rule pruning phase. Each rule constraint is conditionally removed for the rule set and the updated accuracy is tested for improvement.
Finally, the rule set is updated to overcome the inherent overlap that occurs in the rule. To overcome the overlap in each class the range of the attributes for the corresponding classes is modified. For each value of the attribute that is in the overlap range, the probability of its existence in each class is calculated. The value is then shifted to the class having the highest probability. This procedure is repeated for all the attributes.

Final Rule Set:

The final rule set is of the form of IF-THEN propositional clauses. This rule set is used for the classification of the credit risk datasets.

IV. RESULTS

A. 4.1. Data Sets

The experiments are conducted on three commonly used credit-risk evaluation data sets: Australian Credit, German credit and Bank Marketing. All three datasets are publicly available in the UCI repository.
The German dataset consists of 1000 examples with 700 'good' classes and 300 'bad' classes. The dataset has 20 attributes of which 13 are categorical and 7 are numeric. There are no missing values.
The Australian dataset consists of 690 examples with 307 '+' classes and 383 '-' classes. The dataset has 14 attributes of which 8 are categorical and 6 are numerical. 37 examples i.e. 5% have one or more missing values. The mode of the attributes (for categorical) or mean of the attributes (for numerical) are placed in the missing value positions.
The Bank Marketing dataset consists of 4119 examples. There are missing values in a number of examples, so 1000 examples are chosen from the 4119 where each value is present. The sample chosen has 284 'yes' classes and 716 'no' classes. The dataset consists of 16 attributes of which 9 are categorical and 7 numerical.
A description of the three datasets are shown in Table 1.

Table 1. Description of datasets

Dataset	***Attributes***	***Size***
Australian-Credit	14	690
German-Credit	20	1000
Bank-Marketing	16	1000

B. 4.2. Comparison of Results

The performance is evaluated against the following criteria: 10-fold CV accuracy, the average number of antecedents per rules, the Recall, the average False positive rate. 10-fold Cross Validation (CV) is used to make the results more convincing and generalized.
The proposed TDSSCRE is compared with two existing pedagogical models RxNMC[12] and RxREN[16].
From Table. 2 to Table 4. the performance of TDSSCRE against both RxREN and RxNCM is shown.

Table 2. Comparison of results for Australian Credit Dataset

Criteria	***RxREN***	***RxNCM***	***TDSSCRE***
Accuracy	83.64%	85.50%	86.32%
Avg Antec	3	2	2
Racall	0.8333	0.8452	0.8504
FP Rate	0.5666	0.25	0.2333

Table 3. Comparison of results for German Credit Dataset

Criteria	***RxREN***	***RxNCM***	***TDSSCRE***
Accuracy	73.50%	74.67%	75.67%
Avg Antec	2	2	2
Recall	0.7025	0.7233	0.7453
FP Rate	0.5	0.4444	0.4025

Table 4. Comparison of results for Bank Marketing Dataset

Criteria	*RxREN*	*RxNCM*	*TDSSCRE*
Accuracy	84.27%	88.88%	90.26%
Avg Antec	5	3	3
Recall	0.7953	0.8325	0.8543
FP Rate	0.2366	0.1437	0.1466

The rules generated by the two algorithms are described in IF-THEN expression and clauses. The antecedents of the rule with the best performance from all the folds are shown in Tables 5 to 7.

Table 5. Best Rules generated for Australian Dataset

RxREN	*RxNCM*	*TDSSCRE*
if (a8 = 1) then class='+' else class = '-'	if (a10 >=0 and a10<=3) then class='+' else class = '-'	if ((a10 >=0 and a10<=3) and (a14 >=1 and a14<=509)) then class='+' else class = '-'

Table 6. Best Rules generated for German Dataset

RxREN	*RxNCM*	*TDSSCRE*
if (creditHistory = 'existingPaid' and creditAmount <=12204) then class="good" else class="bad"	if ((creditHistory = 'existingPaid' or creditHistory = 'all paid') and creditAmount <=12000) then class="good" else class="bad"	if (creditHistory = 'existingPaid' and creditAmount <=12204) then class="good" else class="bad"

Table 7. Best Rules generated for Bank Marketing Dataset

RxREN	*RxNCM*	*TDSSCRE*
if ((education = 'secondary' or education = 'tertiary') and balance >= 263 and prev_outcome = 'succ') then class = "yes" else class="no"	if ((education = 'secondary' or education = 'tertiary') and (prev_outcome = 'succ' or prev_outcome='unknown')) then class = "yes" else class="no"	if ((education = 'secondary' or education = 'tertiary') and prev_outcome = 'succ') then class = "yes" else class="no"

The above rules correspond to the fold that produced the highest accuracy for the corresponding models.

C. 4.3. Discussion

Performance for Australian Credit Dataset

The TDSSCRE is able to get a higher 10-fold CV accuracy of 86.32 than both RxNCM (85.50%) and RxREN (83.64%). The average number of rule antecedents is 2 antecedents per rule for TDSSCRE, 2 for RxNCM and 3 for RxREN. So, TDSSCRE has a better local comprehensibility than RxREN for having a smaller number of rule antecedents on average. The recall and FP rate of TDSSCRE are better than that of both RxREN and RxNCM. TDSSCRE has a recall of 0.8504 while the recall of RxREN and RxNCM are only 0.8333 and 0.8452 respectively. TDSSCRE has a FP rate of 0.2333 while that of RxREN and RxNCM are 0.5666 and 0.25 respectively. Considering all the measures, TDSSCRE method has overall better performance than that of both RxREN and RxNCM.

Performance for German Credit Dataset

The TDSSCRE is able to get a higher 10-fold CV accuracy of 75.67% than both RxNCM (74.67%) and RxREN (73.50%). The average number of rule antecedents is 2 antecedents per rule for TDSSCRE, RxNCM and RxREN. So, TDSSCRE has an equal local comprehensibility to both RxREN and RxNCM. The recall and FP rate of TDSSCRE are better than that of both RxREN and RxNCM. TDSSCRE has a recall of 0.7453 while the recall of RxREN and RxNCM are only 0.7025 and 0.7233 respectively. TDSSCRE has a FP rate of 0.4025 while that of RxREN and RxNCM are 0.5 and 0.4444 respectively. Considering all the measures, TDSSCRE method has overall better performance than that of both RxREN and RxNCM.

Performance for Bank Marketing Dataset

The TDSSCRE is able to get higher 10-fold CV accuracy of 90.26% than both RxNCM (88.88%) and RxREN (84.27%). The average number of rule antecedents is 3 antecedents per rule for TDSSCRE, 3 for RxNCM and 5 for RxREN. So, TDSSCRE has a better local comprehensibility than RxREN for having fewer number of rule antecedents on average. The recall of TDSSCRE are better than that of both RxREN and RxNCM. The FP rate of TDSSCRE is better than that of RxREN but worse than that of RxNCM. TDSSCRE has a recall of 0.8543 while the recall of RxREN and RxNCM are only 0.7953 and 0.8325 respectively. TDSSCRE has a FP rate of 0.1466 while that of RxREN and RxNCM are 0.2366 and 0.1437 respectively. Taking these into account TDSSCRE method has overall better performance than that of both RxREN and RxNCM except for FP rate when compared to RxNCM.

The results obtained demonstrate that rule extraction using neural network is a viable method for credit risk analysis. The RxNCM algorithm is a suitable choice for this expert system as it has very small computational overhead while producing high accuracy and easily comprehensible rules

V. CONCLUSION

In the ongoing research for developing efficient and explainable credit-risk evaluation models, neural networks have gained a lot of interest recently because of their high level of fault tolerance and property of universal approximation. Works on neural networks primarily focus on improvement of their prediction accuracy without going into how the classifications are being made. This downside of neural network i.e. its lack of explainability is what limits its utilization in credit-risk evaluation domains. In credit-risk evaluation, having concise and comprehensible rules which justify an assessment is essential.

In this paper, a decision system has been devised for credit risk assessment named Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE). The TDSSCRE utilizes the pedagogical rule extraction approach to generate transparent rules. To demonstrate the effectiveness of this algorithm performance comparison has been shown between TDSSCRE and the two pedagogical rule extraction algorithms - RxREN and RxNCM on three real credit-risk evaluation datasets. The comparison is made using: 10-fold CV accuracy, the average number of antecedents per rules, the average recall rate, the average false positive rate. On average the performance of TDSSCRE is greater than that of both RxREN and RxNCM. This makes it suitable to be used as a decision system for credit evaluation.

The propositional rules extracted by the TDSSCRE are concise and comprehensible. Hence the IF-THEN rules are explainable and human understandable to justify the reason for any credit-risk prediction made by the system. Therefore, can be used as a decision support system for credit-risk evaluation in banking sectors without any confusion.

Some other rule extraction algorithms using neural network can be used in future for the credit risk assessment

REFERENCES

[1] V. García, A.L. Marqués, J.S. Sánchez, An insight into the experimental design for credit risk and corporate bankruptcy prediction systems, Journal of Intelligent Information Systems, Volume 44, 2015.

[2] Z. Zhao, S. Xu, B.H. Kang, M.M.J. Kabir, Y. Liu, Investigation and improvement of multi-layer perceptron neural networks for credit scoring, Expert Systems with Applications, 2015.

[3] S.U. Purohit, V. Mahadevan, A.K. Kulkarni, Credit Evaluation Models of Loan Proposals for Indian Banks, International Journal of Modeling and Optimization, Volume 2, 2012.

[4] A.E. Khandani, A.J. Kim, A.W. Lo, Consumer credit-risk models via machine-learning algorithms, Journal of Banking & Finance, Volume 34, 2010.

[5] N. Capon, Credit Scoring Systems: A Critical Analysis, Journal of Marketing, Volume 46, 1982.

[6] H. H. Dam, H. A. Abbass, C. L. and X. Yao, Neural-Based Learning Classifier Systems, IEEE Transactions On Knowledge And Data Engineering, Volume 20, 2008.

[7] R. Caruana and A. Niculescu-Mizil, An Empirical Comparison of Supervised Learning Algorithms, Appearing in Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, 2006.

[8] P.M. Addo, D. Guegan, B. Hassani, Credit Risk Analysis Using Machine and Deep Learning Models, Risks, 2018.

[9] N. Chen, B. Ribeiro, A. Chen, Financial credit risk assessment: a recent review, Artificial Intelligence Review, 2015.

[10] T. Botari, R. Izbicki and A. de Carvalho, Local Interpretation Methods to Machine Learning Using the Domain of the Feature Space, arXiv preprint arXiv:1907.13525, 2019.

[11] R. Setiono and H. Liu, Symbolic representation of neural networks, IEEE Computer, Volume 29, 1996.

[12] S. K. Biswas, M. Chakraborty, B. Purkayastha, P. Roy, D. M. Thounaojam, Rule Extraction from Training Data Using Neural Network, International Journal on Artificial Intelligence Tools, Volume 26, 2017.

[13] E.R. Hruschka and N.F.F. Ebeckenb, Extracting rules from multilayer perceptrons in classification problems: A clustering-based approach, Neurocomputing, Volume 70, 2006.

[14] A.B. Tickle, M. Orlowski and J. Diederich, DEDEC: decision detection by rule extraction from neural networks, QUT NRC, 1994.

[15] K. Jivani, J. Ambasana and S. Kanani, A survey on rule extraction approaches based techniques for data classification using neural network, International Journal of Futuristic Trends in Engineering and Technology, Volume 1, 2014.

[16] M.G. Augasta and T. Kathirvalavakumar, Reverse engineering the neural networks for rule extraction in classification problems, Neural Processing Letters, Volume 35, 2012.