International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 06 | June-2016 www.irjet.net p-ISSN: 2395-0072

A survey on: Predictive Analytics For Credit Risk Assessment

Rahul Vanve[1], Pravin Patil[2]
[1] *Student, Department of computer engineering*
PICT, Pune, India.
[2] *Professor, Department of computer engineering*
PICT, Pune, India.

---***---

Abstract – *In financial domain, credit risk assessment has become important decision process in management activity. Credit risk assessment is a crucial task which helps in a decision process. It can be treated as a classification problem, considering credit risk of the consumer. In such problem aim is to predict the proper classification of the customer into one of the finite class. Based on such classification decision process can be carried out. Credit scoring model can be used to measure the credit risk of a customer. In traditional system most of the research was mainly focused on predictive modelling with statistical approach, with giving less importance to profit maximization. Here such classification is carried out with profit driven approach rather than only statistical approach. Advanced complex non-linear classification techniques are used rather than traditional linear classification technique.*
***Key Words*: Credit Risk, Credit Scoring, Risk Assessment, Predictive Modelling, Classification.**

1. INTRODUCTION

Risk management can be defined as the process of identification analysis and either the acceptance or mitigation of uncertainty in investment decision making. Risk can also be termed as threat [1]. Therefore risk management is about managing uncertainty related to threat. Financial risk management deals with risk that can be managed using traded financial investment. In consumer lending traditional approach is to develop a credit scorecard which rank borrower according to their risk of defaulting [2].

The financial industry's primary task in risk management is to assess the probability of default. In traditional system scoring model with the help of linear discriminant analysis is used for credit risk. Also in some system predictive scorecards have been used for large bank risk management of credit worthiness[1]. Credit scoring models are used in the banking financial sector for the assessment of credit processes. Credit scoring model includes statistical procedures used to classify customers[3]. Consumers who are classified into bad class have a high risk of default and consumers who are classified into good class have a low risk of default. To maximize the profitability of credit card customer, they can be further classified into second classification revolver and transactor. So, credit card customer can be classified with a revolver or transactor scorecard together with a good or bad scorecard which is different from standard approach which deals with only one classification [2].

Credit evaluation is one of the most crucial processes in banks termed as credit management decisions. This process includes collecting, analyzing and classifying different credit elements and variables to assess the credit decisions. Classification of a bankâ€™s customers, as a part of the credit evaluation process to reduce the current and the expected risk of a customer
being bad credit, is credit scoring. Credit scoring can also be termed as, the use of statistical models
to transform relevant data into numerical measures that guide credit decisions [3]. Credit scoring requires less information to make a decision, because credit scoring models have been estimated to include only those variables, which are statistically and significantly correlated with repayment performance. Credit scoring models attempt to correct the bias that would result from considering the repayment histories of only accepted applications and not all applications.

1.1 MOTIVATION

Credit scoring is a very important application in statistical modelling, which distinguishes loan applicant into good or bad class. The main aim is to estimate probability of default of loan applicant. That is the event of a customer not paying back his debt in a given period. Predictive model is developed which assigns a score to each loan applicant. Some threshold value can be set to this score after which customer who have score below this threshold value credit is not granted otherwise accepted. Therefore determining such threshold value is very much crucial for efficient credit risk assessment[5]. One of the most important research issues in financial domain is development of working credit risk evaluation and bankruptcy prediction models. Credit risk is one of frequently faced financial risks, which can be defined as the possibility that counter-party will fail to meet its obligations by agreed terms that will cost invested money for the lender. Minimization of such debts is critical for managing risk in financial institutions [10].

The rapid growth in data capture and computational power has led to an increasing focus on data-driven research. So far, most of the research is focused on predictive modeling using statistical optimization, while profit maximization has been given less priority [7]. In recent years, credit risk assessment has attracted significant attention from managers at financial institutions around the world due to increased demand of consumer credit. These concern is mainly addressed through use of credit scoring model [8]. Many of the statistical methods used to build credit scorecards are based on traditional classification techniques such as logistic regression or discriminant analysis. However, in recent times non-linear approaches, such as the support vector machine and neural networks techniques, have been applied to credit scoring [8].

2. EXISTING SYSTEMS

Credit scoring models are important tools in the credit granting process. These models measure the credit risk of a prospective client. Small and medium sized enterprise can be affected by local economy [6]. In credit risk system probability of default has central role. The models which are used to estimate the probability of default are credit scoring models [6]. In recent years, credit risk assessment has attracted significant attention from managers at financial institutions around the world due to increased demand of consumer credit. These concern is mainly addressed through use of credit scoring model [8]. Many of the statistical methods used to build credit scorecards are based on traditional classification techniques such as logistic regression or discriminant analysis. However, in recent times non-linear approaches, such as the support vector machine, have been applied to credit scoring [8]. Credit scoring is a method of measuring the risk attached to a potential customer, by analyzing their data to determine the likelihood that the prospective borrower will default on a loan. Credit scoring involves various statistical techniques used to convert data into rules which helps organization to take credit granting decisions [8].

Beside Support Vector Machine (SVM), Kernel Logistic Regression (KLR) is one of the most effective and popular methods in the Kernel-machine techniques for classification task in data mining. It is a kernel variant of logistic regression. Logistic Regression has limitation to classify the data with nonlinear boundaries. A non-linear form of Logistic Regression, Kernel Logistic Regression, may overcomes the limitation of Logistic Regression in classifying the data with nonlinear boundaries [13]. Logistic regression model is a multivariate analysis, based on one or more continuous or attribute independent variables, to analyse and predict the attribute dependent variables, mainly used to study the relationship of probability of various dependent variables conditions and the independent variables values. But it can have large mean square error, because of some amount of mutual linear dependence between independent variable [14].

Credit scoring can be described as a classification problem. Traditionally clients have been classified into two groups good and bad [9]. Genetic algorithm is one of the technique which can also be used in the credit scoring technique. Genetic algorithm uses fitness function for classification. The predictive performance of different fitness is evaluated which is used in genetic algorithm [9]. A genetic algorithm uses genetic inspired operations to evolve an initial population into a new population. Here each population comprises of chromosomes that represent genetically encoded individual solutions to a specific problem. Each individual has a fitness score assigned to them, which represents its ability in terms of solution [9]. Credit scoring is mainly a classification problem. In which customers can also be classified into or by profit value based popularity of hybrid technique. In profit based approach, profit is associated with the correct classification and profit may be lost in wrong classification [9].

Transactors are credit card users who pay off their balance every month and so incur no interest charges. Revolver are credit card users who do, occasionally or regularly pay off only part of their monthly balance and so do incur interest charges. Traditional system does not consider such distinction between customer in the decision process of credit granting [2]. The standard approach for building scorecard involves fixed analysis and step-wise regression to identify borrower's characteristics that most affect on customers good or bad status. There are numerous regression, mathematical programming or machine learning techniques have been used for generating scorecard. Here scorecard is generated using logistic regression which is most popular method. Here objective is to present new mechanism of credit and pricing decision scheme[2]. The two main revenue stream from credit card are the merchant service charge and the interest charged on the card balance. The former is a fraction of the value of the purchase made. The latter is charged if the balance is not paid off fully within a given time after a monthly statement being sent [2]. If the transactor/revolver scorecard is used in making the initial accept decision, then it leads to a more exact model of the profitability of the applicants and a more sophisticated acceptance criterion.

Credit risk arises from the very nature of banking business. Understanding credit risk is crucial for the successful functioning of banks. Credit risk as potential losses from the refusal or inability of credit customers to pay what is owed in full and on time. Credit scoring models are developed in order to improve the decision making process within credit risk management. [17] Decision trees are very popular data mining tools. They are used in solving the classification and prediction problems. Decision trees are easy to use, understand and interpret. Business credit scoring models are used to evaluate corporate credit risk. The proper selection of variables is important in model development.

3. CONCLUSION

Large scale default in credit payment by customers reduces profitability and even put the functioning of financial services at risk. It is necessary that credit system take pre-emptive action to avoid or minimize payment default. Predictive model can be used for predicting the likelihood of credit default. The adoption of such strategies can improve customer experience and help improve recovery rate for an organization. Customer can be classified into two classes to grant the credit or not to customer with credit scoring model.

Using advanced classification technique over the traditional classification model improves the performance of credit risk assessment. Considering profit based approach in the classification model, increase the efficiency of credit granting decision for the organization which in turn increase profitability. The performance measure considers benefits generated by healthy loans and help to eliminate the costs caused by loan defaults. As a result, the profit-based approach allows for profit-driven model selection and allows identifying the credit scoring model which increases profitability most. Thus increases profit based credit granting decision.

REFERENCES

[1] Desheng Dash Wu, Shu-Heng Chen, David L. Olson, “Business intelligence in risk management: Some recent progresses”, Information Sciences, vol. 256, no. 20, pp. 1-7, January 2014.

[2] Mee Chi So, Lyn C. Thomas, Hsin-Vonn Seow, Christophe Mues, “Using a transactor/revolver scorecard to make credit and pricing decisions”, Decision Support Systems, vol. 59, no. 1, pp. 143-151, March 2014.

[3] Jos Francisco Martnez Snchez, Gilberto Prez Lechuga, “Assessment of a credit scoring system for popular bank savings and credit”, Contadura y Administracin, Available online January 2016.

[4] Mindy Leow, Jonathan Crook, “Intensity models and transition probabilities for credit card loan delinquencies”, European Journal of Operational Research, vol. 236, no. 2, pp. 685-694, July 2014.

[5] Thomas Verbraken, Cristin Bravo, Richard Weber, Bart Baesens, “Development and application of consumer credit scoring models using profit-based classification measures”, European Journal of Operational Research, vol. 238, no. 2, pp. 505-513, October 2014.

[6] Guilherme Barreto Fernandes, Rinaldo Artes, “Spatial dependence in credit risk and its improvement in credit scoring”, European Journal of Operational Research, vol. 249, no. 2, pp. 517-524, March 2016.

[7] Ashkan Zakaryazad, Ekrem Duman, “A profit-driven Artificial Neural Network (ANN) with applications to fraud detection and direct marketing”, Neurocomputing, vol. 175, no. 1, pp. 121-131, January 2016.

[8] Terry Harris, “Credit scoring using the clustered support vector machine”, Expert Systems with Applications, vol. 42, no. 2, pp. 741-750, February 2015.

[9] Vaclav Kozeny, “Genetic algorithms for credit scoring: Alternative fitness function performance comparison”, Expert Systems with Applications, vol. 42, no. 6, pp. 2998-3004, April 2015.

[10] Paulius Danenas, Gintautas Garsva, “Selection of Support Vector Machines based classifiers for credit risk domain”, Expert Systems with Applications, vol. 42, no. 6, pp. 3194-3204, April 2015.

[11] Maria Rocha Sousa, Joo Gama, Elsio Brando, “A new dynamic modeling framework for credit risk assessment”, Expert Systems with Applications, vol. 45, pp. 341-351, March 2016.

[12] He Ni, “Consumer credit risk evaluation by logistic regression with self-organizing map”, in Natural Computation (ICNC), 2010 Sixth International Conference on, vol. 1, no. 1, pp. 205-209, August 2010.

[13] Rahayu S. P., Zain J. M., Embong A., Purnami S. W., “Credit risk classification using Kernel Logistic Regression with optimal parameter”, in Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on, pp.602-605, May 2010.

[14] Huang Weidong, Zhu Xiangwei, Su Qingling, “Research on application of personal credit scoring based on BPlogistic hybrid algorithm”, in Computer Application and System Modeling (ICCASM), 2010 International Conference on, vol. 4, pp. 735-739, October 2010.

[15] Marin-de-la-Barcena, A. Marcano-Cedeo, A. Jimenez-Trillo, J. Piuela, J. A. Andina, D., “Artificial metaplasticity: An approximation to credit scoring modeling”, in IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, pp.2817-2822, November 2010.

[16] Wang, Xiaoyu Jeong, Dong Chang, Remco Ribarsky, William, "RiskVA: A visual analytics system for consumer credit risk analysis", in Tsinghua Science and Technology, vol. 17, no. 4, pp. 440-451, August 2012.

[17] Kvesic L., Dukic G., "Risk management and business credit scoring", in Information Technology Interfaces (ITI), Proceedings of the ITI 2012 34th International Conference on, pp. 7-52, June 2012.

[18] Hongming Zhou, Yuan Lan, Yeng Chai Soh, Guang-Bin Huang, Rui Zhang, "Credit risk evaluation with extreme learning machine", in Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on , pp. 1064-1069, October 2012.

[19] Jun Huang, Haibo Wang, Wei Wang, Zhibin Xiong, "A Computational Study for Feature Selection on Customer Credit Evaluation", in Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on, pp.2973-2978, October 2013.