Increasing Accuracy of Ensemble Logistics Regression Classifier by Estimating the Newton Raphson Parameter in Credit Scoring

by Armin Lawi

Increasing Accuracy of Ensemble Logistics Regression Classifier by Estimating the Newton Raphson Parameter in Credit Scoring

1st Firman Aziz
fac. mathematics & natural sci.
pancasakti university
Makassar, Indonesia
firmanaziz88@gmail.com

2nd Armin Lawi
department of computer science
hasanuddin university
Makassar, Indonesia
armin@unhas.ac.id
http://orcid.org/0000-0003-1023-6925

3rd Edy Budiman
department of informatics
mulawarman university
Samarinda, Indonesia
edybudiman.unmul@gmail.com
http://orcid.org/0000-0002-3164-5157

***Abstract*— the large volume of customer data in the credit the industry makes the development of an effective credit scoring model extremely important. The use of an ensemble model on statistical methods to solve credit scoring problems managed to get the best predictive performance. Ensemble performance can still be improved by estimating the parameters using nonlinear equations. This paper proposes the estimation of the Ensemble Logistics Regression using Newton Raphson parameter in order to increase the accuracy of the ordinary Logistics Regression model. The results showed that our proposed method successfully improved the accuracy performance of the origin classifier up to 2.6% in the credit scoring.**

Keywords—classifier, ensemble logistic, regression, scoring

I. Introduction

Credit scoring or credit assessment is a valuation method used by banks or other financings, that is useful for determining whether it is feasible or not to get a loan [1]. Credit scoring is a collection of customer data taken from customer loan application data in addition to using a statistical program that contains the history of loans, including how the cycle of payment of bills, payment of bills on time or not, how much credit is still or ever existed. Credit ratings help banks analyse credit applications in addition to other qualitative factors [2]. In addition, with credit scoring creditors can compare debtor information with other customers' loan performance with the same profile.

The function of credit assessment is to help the bank determine whether the loan is approved or not [3], but also determine how much the loan will be obtained, how many conditions are obtained and how much interest the loan will charge. If it turns out the results of a small credit rating, chances are you can still get a loan but with higher interest, or are required to provide collateral. Basically, credit scoring is a matter of classifying the feasibility of prospective debtors [4], that is, candidates for good or bad debtors based on their characteristics.

Lending is based on credit standards. Financial institutions in general already have a selection tool for prospective customers who will be given financing or credit. Credit is given to customers who meet certain specified criteria. However, the selection process does not eliminate the occurrence of poor quality or problematic financing which is popular with the term NPL or non-performing loans.

Supposedly, if the selection or analysis carried out goes well, then the tool for selecting is also good, then the selection results will be obtained that are good, quality or worthy to be given funding. Provision of the appropriate financing to prospective debtors will ultimately improve the quality of financing of the financial institutions concerned. But in reality, debtors who were successfully selected because they were considered worthy in the end there were problems that caused the quality of financing to be bad.

Many method approaches have been developed to try to choose a credit rating model that is better in classifying and predicting the risk of potential debtors.

Classification is a form of data analysis that looks for a set of models, patterns or functions that describe and differentiate data objects to be grouped into specific classes from a number of available classes [5]. In classification, there are three main stages, namely identification or formation of models based on training data, evaluation of models and application of models using data testing. The variables used in the classification consist of predictor variables which are factors that influence or can describe the response variable. In this case, the response variable is in the form of a categorical variable, both of which have order (ordinal) or not (nominal) [6].

Some statistical methods that are often used in classification through categorical data analysis, i.e. Binary Logistic Regression (BLR), Multinomial Logistic Regression (MLR), Ordinal Logistic Regression (OLR), Logistic Regression Ensemble (LORENS) and Log-Linear Regression.

Another method is given in [7] by estimating artificial neural networks and proposing a model in credit ratings. The Logistics Regression model has better performance than other methods on classifying the good and bad debtors. A recent review has been given in [8] to learn about the latest classification algorithms on the application of credit cases. The conclusion stated that for the case of credit score, the Logistics Regression method gets better accuracy results.

This paper proposes the classification of the Ensemble Logistics Regression with the bagging method for credit assessment. the use of the bagging method was chosen to improve the accuracy of artificial and real dataset classifiers by combining a single classifier, and the results were better than random sampling [9]. Bagging is a simple but effective ensemble method and has been applied to many applications in the real world [10]. Bagging can be really useful to build better classifiers when observing training data sets that contain noise [9]. While the focus of this investigation is to estimate the Ensemble Logistic Regression using Newton Raphson parameters. Newton Raphson was chosen because it can solve nonlinear data types by interpreting initial values for their maximum functions.

II. Research Methodology

A. *Credit Scoring*

Credit scoring is a method used to evaluate credit risk in terms of loan applications from consumers [11]. This method is used to classify consumers who apply for credit included in the group of good or bad. Credit scoring attempts to classify the diversity of the characteristics of consumers who submit credit requests based on mistakes and negligence of obligations. This method produces a calculation that can be used by credit service companies to classify the conditions of consumers who apply for credit in relation to credit risk.

To create a "scorecard" scoring model in determining the characteristics of debtors, the development of data analysis is done by looking at the historical data of consumer credit that has been approved by the company or not. The results of this scoring will be useful to predict whether the prospective debtor can carry out the loan well or poorly.

In many cases of scoring systems, with a high scoring value will reduce the value of risk, and the crediting company that provides credit services can determine the calculation limits for accepting or rejecting consumer credit applications based on the value of risk owned. With reference to the credit scoring model that has been formed, the company will approve credit applications if the submitted application has a score above the minimum limit and reject the application if the application submitted has a score below the minimum limit. Although the credit scoring model can further determine the company's policy for accepting or rejecting credit requests from consumers, it is possible to make a prediction error on the value of each potential customer to be given credit facilities. Therefore, to build a good credit scoring model, sufficient historical data is needed.

The credit scoring model is formed through a series of statistical processes that can be used to forecast new data. The process of applying a model that has been formed is different from the process of informing or making a model. In particular, a credit scoring model that is formed can be used for a long time to calculate or predict new data. During the process of establishing the credit scoring model, information from consumers in the form of data is then processed with the help of statistical software. Eventually, a model which has an output in the form of a decision for consumers will be produced.

There are several methods that can be used to produce a credit scoring model, including discriminant analysis, linear regression [12], logistic regression, probability analysis, decision tree [13], and so on. In this research, the method that will be discussed to solve the credit scoring problem is the Ensemble Logistics Regression Classifier method.

B. *Logistics Regression Classification*

The original classification in this study uses Logistic Regression, is one approach to a mathematical model that is used to analyze the relationship of several factors with a dichotomous variable (binary). That is, in binary logistic regression the data on the response variables are binary (0 and 1). These binary numbers describe two opposing data categories, such as good or bad, yes or no, etc.

The equation obtained from the logistic regression process is used to calculate predictions in the credit application process. The bank classifies someone as feasible or not to receive a loan from the bank. Several questions were given to the bank regarding prospective credit recipients. Questions asked about the characteristics of the prospective capital recipient variable are independent variables that will be inputted by bank officers into the model. From some of the variables in question, bank officers can determine the chances of the prospective recipient of the loan being able to return the loan or not, a value between 0-1.

Scores of the model use binary numbers where 1 for good debtors and 2 for bad debtors, and each independent variable gives a dependency on each score [14]. The classifier function is shown by "equation (1)".

$$\log\left(\frac{p}{1-p}\right) = \sum_{i=1}^{i=n} \beta^{(i)} * x^{(i)} + e = \beta^T x + e, \qquad (1)$$

Where β is a parameter, e an error i, where i = 1, 2, ..., n.

Logistic regression will form predictor or response variables, which is a linear combination of independent variables. The value of this predictor variable is then transformed into a probability by the logit function.

The logistic model is expressed in the form of a probability model in which this response variable is the logit of the probability of a situation or attribute that will apply with the terms or conditions of certain independent variables. Following this is the logistical regression probability model found in "equation (2)".

$$p = \frac{e^{\beta^T x + e}}{1 + e^{\beta^T x + e}} \qquad (2)$$

The estimation of Logistic Regression parameters can be obtained in two ways, i.e., Maximum Likelihood Estimation (MLE) and Newton Raphson. However, this study uses the parameters of Newton Raphson. Newton Raphson is a method for solving nonlinear equations such as solving likelihood equations in the Logistics Regression models (14). The Newton Raphson method requires an initial estimation of its maximum function value, in which the estimation using a polynomial approach of degree two.

In this case, to determine the estimation of β value of β which is the maximum function of $g(\beta)$. Suppose $q' = \left(\frac{\partial g}{\delta\beta 1}, \frac{\partial g}{\delta\beta 2}, \ldots\right)$, and suppose H is denoted as a matrix with members $h_{ab} = \frac{\partial g}{\delta\beta 1, \delta\beta 2}$ suppose $q^{(t)}$ and $\boldsymbol{H}^{(t)}$ is a form of evaluation of $\beta^{(t)}$ estimate to t at β. In step t in the iteration process (t = 0, 1, 2, ...).

C. Ensemble Bagging Classification

The focus of this classification is to solve similar problems by combining a set of classifications to obtain a more accurate classification [15], [16]. Bagging is a method that combines bootstrapping and aggregating [9].This bootstrap sample is obtained by resampling with replacements from the original dataset to produce the same number of elements from the original dataset. Bagging can be really useful to build better classifiers when observing training data sets that contain noise. The main idea of the ensemble method is to combine several sets of models that solve a similar problem to get a more accurate model. Illustration of the ensemble method can be seen in "Figure 1".

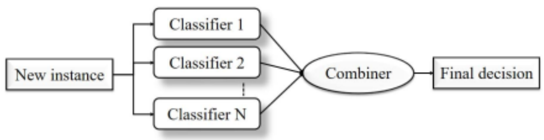

Fig.1. Illustration of ensemble method [17]

Bagging is good for classification and regression [18]. In the case of regression, to be better, one can take averages when combining predictions. Bagging one of the stable learning algorithms on small changes in the training set causes a large difference in the students produced, namely the learning algorithm on data that has high variance (noise). High noise affects the new classifications produced, thus causing miss-classification [19]. Noise in terms of prediction accuracy can be improved by using classifier-ensembles [20].

Bagging can significantly increase accuracy significantly compared to individual models and is stronger against the effects of noise and overfitting than the original training data [5], [21]. Bagging Algorithm [18]:

Looping for b = 1, 2,. . ., B

- Create a bootstrap sample $\{(X, Y1 *)\ 1 *, (X, Y2 *), \ldots, 2 * (Xn *, Yn *)\}$ with random replacement from training data $\{(X1, Y1), (X2\ , Y2), \ldots, (Xn, Yn)\}$ matches the Cb classifier turned on in the appropriate bootstrap sample.
- Final classifier output calculated using "equation (3)".

$$C(x) = B^{-1} \sum_{b=1}^{B} C_b(x) \tag{3}$$

D. Performance Evaluation

To measure performance, use confusion matrix. Confusion matrix provides decisions obtained in training and testing [22]. Confusion matrix provides an assessment of the performance of classifications based on true or false objects [23]. Confusion matrix is a 2-dimensional matrix that illustrates the comparison between predicted results and reality [24]. For unbalanced data, accuracy is more dominated by accuracy in minority class data, then the right metrics are AUC (Area Under the ROC Curve), F-Measure, GMean, overall accuracy, and accuracy for minority classes [25]. Minority class accuracy can use the TPrate / recall (sensitivity) metric. G-Mean and AUC are more comprehensive evaluators of predictors in the context of imbalance [26]. To do the calculations used "equation (4)" [23]:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

III. Results and Discussion

The focus of this research is to look at the use of Newton Raphson parameters from the Logistic Regression classification ensemble. This is because the estimation of Newton Raphson parameters can improve the performance of the Logistic Regression classification ensemble. Logistic Regression will be used as a single classification and the ensemble technique used is bagging.

A. Dataset

The data used and processed to test the proposed model are Australian and German datasets. Both datasets are used to test the classification model for credit rating problems. General information is presented in Table I.

TABLE I. Description of Dataset

Dataset	#Debtors	Debtors (good)	Debtors (bad)	#Attributes
australian	690	307	383	14
german	1,000	700	300	20

Next, the dataset in Table I compares 70% (training) and 30% (testing) to assess model performance.

B. Performance Evaluation

Performance evaluation of the proposed non-linear ensemble Logistics Regression classification method uses test processes written in the python programming language. The level of accuracy based on our application results is given in Table II both for the ordinary Logistics Regression model and the model which is resulted from the estimation of Newton Raphson parameter of the Ensemble Logistics Regression.

TABLE II. Classification Results

Classification	German dataset (%)	Australian dataset (%)
Logistic regression	77.0	85.9
Ensemble logistic regression	79.6	86.9

Table II shows the classification results using ordinary Logistic Regression model for the German and Australian datasets, respectively, gives classification accuracy of 77.0% and 85.9%. Whereas, the Ensemble Logistics Regression based on the estimation of Newton Raphson parameter increases the accuracy to 79.6% and 86.9% for both German and Australian datasets, respectively.

Although the increasing accuracy is not very significant due to the Australian dataset only increased by 1% and the German dataset only increased by 2.6%. but in this study obtained when the number of iterations 10 to 50 affects the accuracy of the ensemble classification Logistic Regression but if iteration> 50 accuracy value becomes constant. the proposed method is sensitive to small amounts of data.

IV. Conclusion

This paper has proposed an ensemble Logistics Regression method with bagging technique for credit scoring. The focus of the investigation is to estimate the ensemble Logistics Regression model using the Newton Raphson parameter. The proposed method successfully increased the accuracy of the ordinary Logistics Regression classification. The accuracy method yields 79.6% and 86.9% for German and Australian dataset, respectively. The number of iterations 10 to 50 affects the accuracy of the classification, however, constant accuracy value has happened when the iteration > 50 and the proposed method is sensitive to the small amount of data.

References

[1] F. Pandia, "Manajemen Dana dan Kesehatan Bank," in *Manajemen Dana dan Kesehatan Bank*, 2012.

[2] L. C. Thomas, D. B. Edelman, and J. N. Crook, "Credit scoring and its applications," *SIAM Monographs on Mathematical Modeling and Computation*, 2002.

[3] M. Modina and M. Modina, "Credit Risk Management," in *Credit Rating and Bank-Firm Relationships*, 2015.

[4] H. M. Ritonga, H. A. Hasibuan, and A. P. U. Siahaan, "Credit Assessment in Determining The Feasibility of Debtors Using Profile Matching," *International Journal of Business and Management Invention*, 2017.

[5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Third Edit. California, 2012.

[6] M. Wati, N. Puspitasari, E. Budiman, and Haeruddin, "Multi-Criteria Decision-Making for Evaluation of Student Academic Performance Based on Objective Weights," in *2018 Third International Conference on Informatics and Computing (ICIC)*, 2018, pp. 1–5.

[7] V. S. Desai, J. N. Crook, and G. A. Overstreet, "A comparison of neural networks and linear scoring models in the credit union environment," *European Journal of Operational Research*, 1996.

[8] S. Lessmann, B. Baesens, H. V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *European Journal of Operational Research*, 2015.

[9] E. Alfaro, M. Gáamez, and N. García, "Adabag: An R package for classification with boosting and bagging," *Journal of Statistical Software*, 2013.

[10] G. Liang and C. Zhang, "Empirical study of bagging predictors on medical data," *Conferences in Research and Practice in Information Technology Series*, 2010.

[11] W. H. Greene, "A statistical model for credit scoring," in *Advances in Credit Risk Modelling and Corporate Bankruptcy Prediction*, 2008.

[12] E. Budiman, Y. Parassa, H. Haerullah, D. Moeis, R. Soekarta, and M. Jamil, "The effect of mobile learning media on student learning motivation in data structure courses," *Proceedings - 2018 3rd International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2018*, pp. 60–64, 2019.

[13] E. Budiman, N. Dengen, Haviluddin, and W. Indrawan, "Integrated multi criteria decision making for a destitute problem," in *2017 3rd International Conference on Science in Information Technology (ICSITech)*, 2017, vol. 2018-Janua, pp. 342–347.

[14] L. C. Thomas, "A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers," *International Journal of Forecasting*, 2000.

[15] A. Lawi, F. Aziz, and S. Syarif, "Ensemble GradientBoost for increasing classification accuracy of credit scoring," in *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology, CAIPT 2017*, 2018.

[16] L. Zhou, K. K. Lai, and L. Yu, "Least squares support vector machines ensemble models for credit scoring," *Expert Systems with Applications*, 2010.

[17] G. Wang and J. Ma, "A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine," *Expert Systems with Applications*, 2012.

[18] L. Breiman, "Bagging predictors - Springer," *Machine learning*, 1996.

[19] D. Kim, P. Kang, S. Cho, H. J. Lee, and S. Doh, "Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing," *Expert Systems with Applications*, 2012.

[20] B. Twala, "Multiple classifier application to credit risk assessment," *Expert Systems with Applications*, 2010.

[21] M. Culp, G. Michailidis, and K. Johnson, "On adaptive regularization methods in boosting," *Journal of Computational and Graphical Statistics*, 2011.

[22] E. Budiman, U. Haryaka, J. R. J. R. Watulingas, and F. Alameka, "Performance rate for implementation of mobile learning in network," in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017, vol. 2017-Decem, pp. 1–6.

[23] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intelligent Systems Reference Library*, 2011.

[24] E. Budiman, Haviluddin, N. Dengan, A. H. A. H. Kridalaksana, M. Wati, and Purnawansyah, "Performance of Decision Tree C4.5 Algorithm in Student Academic Evaluation," in *Lecture Notes in Electrical Engineering*, vol. 488, R. Alfred, H. Iida, A. A. Ag. Ibrahim, and Y. Lim, Eds. Singapore: Springer Singapore, 2018, pp. 380–389.

[25] H. Zhang and Z. Wang, "A normal distribution-based over-sampling approach to imbalanced data classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011.

[26] S. Wang and X. Yao, "Using class imbalance learning for software defect prediction," *IEEE Transactions on Reliability*, 2013.

Increasing Accuracy of Ensemble Logistics Regression Classifier by Estimating the Newton Raphson Parameter in Credit Scoring

ORIGINALITY REPORT

10%	3%	9%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Basrah - College of Science Student Paper	2%
2	T Desyani, Y Kasmayanti, A Saifudin, Yulianti. "Bagging Techniques to Reduce Misclassification of Breast Cancer Prediction Base on Gradient Boosted Trees (GBT) Algorithm", Journal of Physics: Conference Series, 2020 Publication	2%
3	Dina Fitrisia Septiarini, et al.. "Analytical Hierarchy Process (AHP) Technique in Productive Financing Decision Making(A Case Study on Islamic Banks in East Java,Indonesia)", Journal of Islamic Financial Studies, 2016 Publication	1%
4	Hernawati Pramesti, Herman R.A. Talompo. "Determination of Credit Decision Attributes Using Maximal Information Coefficient", 2018	1%

International Conference on Information Technology Systems and Innovation (ICITSI), 2018

Publication

5	E Sugiharti, A T Putra, Subhan. "Facial recognition using two-dimensional principal component analysis and k-nearest neighbor: a case analysis of facial images", Journal of Physics: Conference Series, 2020 Publication	1%
6	Ranran Li, Chongchong Zhao, Xin Li, Guigang Zhang, Yong Zhang, Chunxiao Xing. "Chapter 29 Comparative Analysis of Medical P2P for Credit Scores", Springer Science and Business Media LLC, 2018 Publication	1%
7	A Saifudin, S W H L Hendric, B Soewito, F L Gaol, E Abdurachman, Y Heryadi. "Tackling Imbalanced Class on Cross-Project Defect Prediction Using Ensemble SMOTE", IOP Conference Series: Materials Science and Engineering, 2019 Publication	1%
8	www.ojs.excelingtech.co.uk Internet Source	1%