Communication in Mathematical Modeling and Applications
http://ntmsci.com/cmma

An application of logistics regression model to determining the credit suitability and impacting factors in a special bank branch

Muharrem Unver, Berge Sahin and Filiz Ersoz

Department of Industrial Engineering, Engineering Faculty, Karabuk University, Karabuk, Turkey

Received: 24 August 2017, Accepted: 15 January 2018
Published online: 26 January 2018.

Abstract: There are quite complicated rules and constraints that can be imposed by the bank when the loan issued. Bank branches, which play a direct role in the credit, must accurately determine the customer's credit request to eliminate these difficulties and create an effective payment system according to the customer. In the study, 100 random loan applications made in 2016 of a bank branch operating in the Black Sea Region were examined. These customer demands are affecting customer characteristics. The "Logistic Regression (LR) Model" was created to predict creditworthiness according to the identified fugitives. In the model, customer age, education, marital status, debt grade, credit card debt, other debts, cross product are the variables. These are statistically significant in terms of marital status, gender, cross product, or creditworthiness. However, various variables such as debt income ratio, credit card debt, and other debts are statistically significant and affect credibility to negatively. In addition, occupational, income and educational constraints were found to be meaningless. With this model, the factors affecting the credit were evaluated. As a result of the study, the bank branch will benefit from the statistical model in which it is created, to evaluate according to the customer characteristics in its portfolio, and to give more credit to branch customers.

Keywords: Bank, logistic regression (LR), factors affecting for loan.

1 Introduction

If people are not enough to obtain the financial means they need, they demand it in various forms. In the last quarter of the 20th century, the use of credit by the financial sector in Turkey started to evolve very rapidly, with the introduction of various financial institutions outside the bank, as well as public and private banks. Nowadays, there are a large number of institutions in the Turkish financial system that are becoming brands or specializing in areas such as "insurance companies", "mutual funds", "partnerships", "consulting companies" and "leasing". But, there are quite complicated rules and constraints that must be made by the bank when credits are given to these institutions. Bank branches that play a direct role in the credit should correctly determine the customer's credit request to overcome these difficulties and establish an efficient payment system according to the customer.

The credit risk problem is substantially the computation of the loss level, which is described as the level for which there is a possibility of 1% that the loss incurred in the portfolio will overrun that level in a specific time period. Credit risk has been the subject of much research activity, especially after realizing its practical necessity after a number of high profile bank failures in Asia. As a result, the regulators are acknowledging the need and are urging the banks to utilize cutting-edge technology to assess the credit risk in their portfolios. Measuring the credit risk exactly withal allows banks to engineer future loan transactions, so as to achieve targeted return and risk characteristics [1].

* Corresponding author e-mail: muharremunver@karabuk.edu.tr

In the study, 100 random loan applications made in 2016 of a bank branch operating in the Black Sea region were examined. In this context, the effect on credit characteristics of customers applying for credit has been investigated. In this paper, a literature search was conducted first. Then, ”Logistic Regression Analysis” was conducted to determine the problem and related factors and to predict the credibility according to these factors. In the model, customer age, education status, marital status, gender, profession, income, debt income ratio, credit card debt, other debts and multiplication product are taken as independent variables. So, the credibility determined based on customer characteristics; A regression model was set up to answer the question of whether or not the loan should approved.

2 Literature review

Maher et al. in their study, utilizes a neural network approach to modeling the bond rating process in an attempt to increase the overall prediction accuracy of the models. A comparison is made to a more traditional LR approach to classification prediction. The results indicate that the neural networks based model performs significantly better than the LR model for classifying a holdout sample of newly issued bonds in the 1990–92 period. A potential drawback to a neural network approach is a tendency to overfit the data which could negatively affect the model’s generalizability. In their paper, carefully controls for overfitting and obtains significant improvement in bond rating prediction compared to the LR approach [2].

In their study, Baesens et al., they study to performance of various state of the art classification algorithms applied to eight real life credit scoring data sets. Some of the data sets originate from Benelux and England financial institutions. The performance is assessed using the classification accuracy and the area under the receiver operating characteristic curve. Statistically significant performance differences are identified using the appropriate test statistics. It is found that both the LS-SVM and neural network classifiers yield a very good performance, but also simple classifiers such as LR and discriminant analysis perform very well for credit scoring [3].

Zekic et al. in their study, compares the models for small business credit scoring developed by LR, neural networks, and CART decision trees on a Croatian bank dataset. The models obtained by all three methodologies were estimated; then validated on the same hold-out sample, and their performance is compared. There is an evident significant difference among the best neural network model, decision tree model, and LR model. The most successful neural network model was obtained by the probabilistic algorithm. The best model extracted the most important features for small business credit scoring from the observed all data [4].

Bensic et all., the main purpose of their study is to extract important features for credit scoring in small-business lending on a dataset with specific transitional economic conditions using a relatively small dataset. To do this, they compare the accuracy of the best models extracted by different methodologies, such as LR, neural networks (NN), and CART decision trees. Four different NN algorithms are tested, including backpropagation, radial basis function network, probabilistic and learning vector quantization, by using the forward nonlinear variable selection strategy. Although the test of differences in proportion and McNemar’s test do not show a statistically significant difference in the models tested, the probabilistic NN model produces the highest hit rate and the lowest type I error. According to the measures of association, the best NN model also shows the highest degree of association with the data, and it yields the lowest total relative cost of misclassification for all scenarios examined. The best model extracts a set of important features for small-business credit scoring for the observed sample, emphasizing credit programme characteristics, as well as entrepreneur’s personal and business characteristics as the most important ones [5].

Su-Juan in his study, he created a LR credit appraisal model using Spss. It has used 106 listed companies in China to separate two patterns. The two patterns mean that the listed companies are divided into two groups according to the business conditions: one is a ”bad” group and the other is a ”normal” group. For each listed company, four main financial issues are considered: earnings per share, net asset per share, return on equity, cash flow per share. The simulation results show that the differential accuracy rate is 99.06% using the LR credit evaluation model, according to 106 example. As a result, the distinctive results of the LR model are not as good as the discriminant analysis [6].

Lee et al. in their study, explore the performance of credit scoring using two commonly discussed data mining techniques classification and regression tree (CART) and multivariate adaptive regression splines (MARS). To demonstrate the effectiveness of credit scoring using CART and MARS, credit scoring tasks are performed on one bank credit card data set. As the results reveal, CART and MARS outperform traditional discriminant analysis, LR, neural networks, and support vector machine (SVM) approaches in terms of credit scoring accuracy and hence provide efficient alternatives in implementing credit scoring tasks. In order to evaluate the feasibility and effectiveness of using CART and MARS in building credit scoring models, credit scoring task is performed on one bank credit card data set. Analytic results demonstrate that CART and MARS both have better average correct classification rate in comparison with discriminant analysis, LR, neural networks, and support vector machine [7].

Ata in his study, foreign banks in the Turkish banking sector are being compared in terms of their financial performance and analyzed the effect of the sectoral foreign capital inflow on the performance of the banks. In this study, they performed a highly variable LR analysis in terms of profitability, efficiency, liquidity and risk factors in terms of domestic and foreign banks. In the analysis, domestic and foreign banks operating in Turkey used the data for 2002-2007 period. As a result, in terms of performance indicators, domestic banks were found to be more effective than foreign banks [8].

Dong et al. in their study, have proposed a LR model with random coefficients for the construction of credit scorecards. Among them, the LR model is the most widely used model in the banking sector because it has best preferred features. Although some new techniques have been applied to credit ratings and show no superior predictive accuracy, problems arise with the interpretability of the results. A LR model with random coefficients is proposed to improve the accuracy of the prediction of the LR. The proposed model can improve the accuracy of forecasting the LR without compromising the desired characteristics. It is expected that the proposed method of generating credit score cards may contribute to the effective management of credit risk in practice. Empirical results show that the proposed model can improve the predictive accuracy of the LR with constant coefficients without compromising the desired characteristics [9].

Zhu and Li have studied this issue as investors began to worry about analyzing credit risk for listed companies, with the emergence of stock market credit issues and the frequent credit crisis. Taking into account the current development methods of the credit risk analysis and the importance of determining the corporate financial risk, an effective indicator system has been designed in their study and the credit evaluation models of China’s stock market companies have been established by taking advantage of the 2009 financial data. When China merged with the facts of the companies traded on the stock exchange, they used established models to discriminate and analyze. The outcome of the empirical research on credit risk analysis for firms traded on the stock market is superior to the differential analysis model of the LR model [10].

Nie et al. in their study, two data mining algorithms are applied to build a churn prediction model using credit card data collected from a real Chinese bank. The contribution of four variable categories: customer information, card information, risk information, and transaction activity information are examined. In their paper, analyzes a process of dealing with variables when data is obtained from a database instead of a survey. Instead of considering the all 135 variables into the model directly, it selects the certain variables from the perspective of not only correlation but also economic sense. In addition to the accuracy of analytic results, the study designs a misclassification cost measurement by taking the two types error and the economic sense into account, which is more suitable to evaluate the credit card churn prediction model. The algorithms used in their study include LR and decision tree which are proven mature and powerful classification algorithms. The test result showed that regression performs a little better than decision tree [11].

Budak and Erpolat in their study, artificial neural networks and LR analysis have been used to provide a support to the Banks' credit risk prediction and to estimate whether a credit demanding customers' repayment order will be regular or not. The results of the study showed that artificial neural networks method is more reliable than LR analysis [12].

Kara and Kuru in their study, the way in which the consumer's efforts in Yozgat's banking practices affected the consumers, the demographic characteristics of the consumers, as well as making suggestions about the efforts to be made for the banks. They obtained the data of the survey by questionnaire. The percentages and frequencies of the obtained data were taken and it was determined whether the bank's sales activities affected the purchasing decision by considering the demographic characteristics of the consumer [13].

Bekhet and Eletter in their study, presented two models for credit scoring applications in commercial banks in Jordan: RBF (radial basis function) and LR. The loan application review increased the effectiveness of the credit decision and controlled the credit bureau duties, reducing the analysis time and cost. Acceptance and rejection of credit applications from other commercial banks were used to create credit scoring models. The results show that RBF is more accurate and interpretive than the LR model, although it does show that it induces encouraging results for screening of bad practices. As a result, the LR model appears to perform slightly better than the radial-based function model in terms of general accuracy. However, the radialbased function is superior to determining which customers may be assumed. Both models showed promising results and concluded that there was no general best model for loan application evaluation. LR gave better results than RBF in terms of overall classification ratio [14].

Khemais et al. in their study, aimed to develop a modeler to anticipate the default risk of small and medium-sized businesses used for two different methodologies for Tunisia's one of the trade bank. They used a database that consists of 195 credit files granted to Tunisian SMEs which are divided into five sectors "industry, agriculture, tourism, trade and services" for a period from 2012 to 2014. The empirical results that they found support the idea that these two scoring techniques have a statistically significant power in predicting default risk of enterprises. Logistic discrimination classifies enterprises correctly in their original groups with a rate of 76.7% against 76.4% in case of linear discrimination giving so a slight superiority to the first method [15].

3 Material and method

In line with the data used in the study, 100 random loan application made in 2016 was handled by a private bank operating in our country in line with customer credit information in a branch in the Black Sea region. In the study, it was researched whether the information about the age, gender, occupation, marital status, income, debt income rate and education of the client has a credibility rating. So, it was aimed to provide more effective service in determining the new customer portfolio characteristics that will be formed by looking at the customer portfolio of the bank branch in terms of determining more accurately according to the credit or affecting the feature development opportunities.

In this study "Binary Lojistic Regression" model is used. Multiple independent variables and a dependent variable are used in the model. The dependent variable is set to yes-no (0-1) to answer the question "Is the credit approved?" The independent variables are; Age, gender, occupation, marital status, education, income, debt income ratio, and duration of work with the bank (year). The data obtained from the center was processed into the "SPSS 23.0" program and the "LR Model" analysis was applied. In the analysis, factors that are effective after credit approval or rejection and the impact ratings of these factors are investigated. The logistic distribution function that can explain the model is as follows:

$$\rho_i = \varepsilon(\gamma_i = 1/X_i) = \frac{1}{1+e^{-(\beta_1+\beta_2 X_i \ldots)}} \tag{1}$$

It is also possible to use intermittent, continuous and categorical data in the LR. In this manner, this model emerges as a strong and more descriptive regression model. In LR analysis, some of the variables studied were used to show cause-effect relationships; Yes-no positive-negative, satisfied-not satisfied 2-level data. ”LR Analysis” is important when the dependent variable is composed of 2 or more levels of categorical data and the causal relation between dependent variable and independent variables is examined [16].

$\phi(x)$ in equation, also satisfies the constraint which bond of probability Y given X should be between 0 and 1. One important study in LR is the *logit transformation* where *“odds”* are introduced, i.e. $\phi(x)1-\phi(x)$ means the probability of an event relative to the probability of the event not happening. The logit transformation is given in terms of $\phi(x)$:

$$g(x) = ln((\pi(x)/(1-\pi(x)) = \mu + \beta 1x1 + \beta 2x2 + .. + BJxJ = \beta \times 1\mu + x'\beta. \tag{2}$$

When **x** increases one unit, the odds ratio increases $exp(\beta)$ unit. [17]

4 Data set and determination of variables

Findings dealt with in the study are described as dependent variables and independent variables as follows. The dependent variable is categorical and has two outcomes and represents (0,1) yes and no separately. Some of the independent variables were evaluated on the nominal scale and some on the ordinal scale scale. At the time of the study, 100 randomly selected credit applications were examined in 2016 and the credit / non-withdrawal status; ”0” credits are coded as ”1” credits.

Independent Variables:

–Age: Applicant’s age
–Education: Level of the applicant’s education
–Year: Years worked
–Income: Household income
–Debt_Income: Debt ratio to income
–Cross_product: The number of customer’s cross_product.

The dependent Variables:
Credits:

–Credit = 1: Bank approves credit
–Credit = 0: The bank has not approved the credit

is defined as. Descriptive statistics were calculated to reveal the general characteristics of the data used in the study and are presented in Table 1 below.

Table 1: Descriptive statistics

	Age	Edu.	Year Worked	Houseolds Income	Debt_ Income Ratio	Credit Approval	Cross_ Product Number
N Valid	100	100	100	100	100	100	100
Missing	0	0	0	0	0	0	0
Mean	39,13	2,89	8,39	4053,600	10,1520	,69	3,7800
Std. Error of Mean	1,404	0,114	,676	263,21760	,71694	,046	,23076
Std. Deviation	14,044	1,136	6,755	2632,1760	7,16943	,465	2,30756
Variance	197,246	1,291	45,634	6928350,5	51,401	,216	5,325
Minimum	18	1	0	1260,00	,40	0	,00
Maximum	85	4	26	15840,00	41,30	1	9,00

Information indicating the educational level, marital status, gender and credit approval status of the applicant is given in Table 2 ”Frequency Table”.

When Table 2 is examined, 31% bank not approve loan, 69% approve loan. When the education levels of the customers are examined, 18% is primary school, 16% is middle school, 25% is high school and 41% is university graduate. As for sex, 59% are male and 41% are female. Customer’s marital status; 60% are married, 28% are single, 12% are widow. When looked at cross products; 4% of customers without any product. 1 product, 9%. 2 products, 24%. 3 products, 14%. 4 products, 16%. 5 products, 12%. 6 products, 7%. 7 products, 2%. 8 products, 9%. 9 products were found as 3%.

Table 2: Frequency table.

Research	Status	Frequency
What is the level of education of the customer?	Primary School	18
	Middle School	16
	High School	25
	University and/or postgraduate	41
	Total	100
What is the customer’s marital status?	Married	60
	Single	28
	Widow	12
	Total	100
What is the customer’s gender?	Man	59
	Woman	41
	Total	100
Is the credit approved?	Yes	69
	No	31
	Total	100
How many the cross product?	Zero	4
	One	9
	Two	24
	Three	14
	Four	16
	Five	12
	Six	7
	Seven	2
	Eight	9
	Nine	3
	Total	100

The Enter Method has been implemented in the course of measuring the effect of variables (credibility). The ”β” parameters and the ”Wald” statistical values, the degree of freedom, the significance levels and the ODDS (Odds Ratio) of this effect were examined in the above table.

Table 3: Significance levels of variables affecting credit.

		Score	df	Sig.
Step 0 Variables	**Age**	3,005	1	0,083
	Education	3,418	3	,332
	Education(1)	1,855	1	,173
	Education(2)	,376	1	,540
	Education(3)	,016	1	,901
	Year	4,658	1	,031
	Income	4,715	1	,030
	Debt_Income	33,232	1	,000
	Marital_Status	16,105	2	,000
	Marital_Status(1)	8,485	1	,004
	Marital_Status(2)	16,053	1	,000
	Gender(1)	3,557	1	,059
	Cross_Product	18,913	1	,000

Significance of the characteristics affecting credibility according to the level of importance according to the results of regression; The duration of the client's work with the bank, the debt income ratio (Debt_Income), the marital status of the client (Marital_Status), gender and the number of cross products were found to be significant. The results obtained when the variables affecting creditworthiness are investigated by the LR model are given in Table 4 below.

Table 4: Logistics Regression Results

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1a	**Age**	-0,025	,049	,255	1	,614	,975
	Education	0	0	,901	3	,825	0
	Education(1)	-,724	1,784	,165	1	,685	,485
	Education(2)	,703	1,754	,161	1	,688	2,020
	Education(3)	,688	1,253	,301	1	,583	1,989
	Year	,387	,155	6,243	1	,012	1,473
	Income	,000	,000	1,652	1	,199	1,000
	Debt_Income	-,596	,169	12,500	1	,000	,551
	Marital_Status	0	0	6,222	2	,045	0
	Marital_Status(1)	-2,325	1,631	2,032	1	,154	,098
	Marital_Status(2)	-4,673	1,934	5,840	1	,016	,009
	Gender(1)	2,061	1,002	4,234	1	,040	7,853
	Cross_Product	1,199	,382	9,866	1	,002	3,316
	Constant	5,398	2,876	3,521	1	,061	220,859

$$\begin{aligned} \text{Y} =& 5,398+0,387\,(\text{Year})-0,596\,(\text{Debt\_Income})-4,673\,(\text{Marital\_Status}\,(2)) \\ &+2,061\,(\text{Gender}\,(1))+1,199\,(\text{Cross\_Product}) \end{aligned} \tag{3}$$

The logistics regression distribution function is given below.

$$P=\frac{1}{1+e^{-(5{,}398-0{,}25X_1-0{,}724X_2+0{,}703X_3+0{,}688X_4+0{,}387X_5+0X_6-0{,}596X_7-2{,}325X_8-4{,}673X_9+2{,}061X_{10}+1{,}199X_{11})}} \tag{4}$$

According to the results of LR; The most important variables in deciding whether or not to give credit to the bank's customers are as follows: Marital status, gender, cross-product and educational change. The number of cross-products of the customer is expected to increase by 3,316 times compared to the small number. The higher the cross-product of the client's credit exposure, the more positive the factor will be at 5,398.

Gender (1) has become a result of the fact that if the male is a male, the situation will be increased by 7,853 times, and the credit affects positively. In addition, if the sex is male, the credit outcome will affect positively with a factor of 2,061. When the client is looking at the year in which he / she works with the bank; This increase in employment will affect the credit outlook by 1,473 times positively. This is due to the fact that the working time with the bank is high and the credit outcome will have a positive impact on the factor of 0.387.

When the ratio of the client's Debt_Income is investigated according to the borrowing rate, the higher the factor is, the more difficult it is for the borrower to make a loan. The effect of this situation is 0.551 times more. The high debt to income ratio is assessed with a factor of -0.596, affecting the credit outcome in the negative direction.

Since the study included general information about the customers in this study, some criteria were not found to be significant according to credibility. This situation was found to be normal only to the customers because of the question of whether the loan was issued or not. In particular, it has been seen that customer education and income do not affect creditworthiness very much.

In the interpretation of the model obtained as a result of LR Analysis, the suitability of the model should be tested and its results should be examined. According to the model classification table, in the credibility model, the probability of correctly estimating the non-credit status of a customer who has not a loan is 83.9% as a result of the characteristics of the customer after the credit and other effects of the credit. The rate of accurately estimating a loan is 95.7%. The correct classification rate of the model is 92%.

Table 5: Model classification table.

Observed			Predicted		
			Credit Approval		**Percentage Correct**
			No	Yes	
Step 1	**Credit Approval**	No	26	5	83,9
		Yes	3	66	95,7
	Overall Percentage				92,0

When the significance of the constructed LR model is evaluated, it can be said that the values found meaningful according to the Wald test statistic are significant by looking at the values of 1% significance level (sig.).

The Omnibus Test is seen in Table; when the model with significance level of ”0,01” is significant. According to the table, the degree of freedom was calculated as 11, and the Chi- square value as 86,328. The other test used to test the fit of the model is the ”Hosmer and Lemeshow” test. It tests whether all logit coefficients outside the fixed term are zero.

Table 6: Omnibus test of model coefficients.

		Chi-square	df	Sig.
Step 1	Step	86,328	11	,000
	Block	86,328	11	,000
	Model	86,328	11	,000

The established hypothesis for model coefficients is given below.
H0: Parameters have a good discrimination in terms of decisiveness
Hs: Parameters do not have good discrimination in terms of decisiveness
Chi square values for the Hosmer-Lemeshow test statistic for hypotheses:

Table 7: Hosmer-Lemeshow test.

Step	Chi-square	df	Sig.
1	1,227	8	,996

According to Hosmer-Lemeshow test result; model; The H0 hypothesis was accepted in the case of 8 degrees of freedom with a value of 1,227 Chi-square, p = 0,996> 0,05. That is, the parameters have a very good discrimination in terms of determinism.

Table 8: Observed and expected frequencies of the Hosmer-Lemeshow test.

		Credit Approval=No		Credit Approval=Yes		
		Observed	Expected	Observed	Expected	Total
Step 1	1	10	9,989	0	,011	10
	2	9	9,448	1	,552	10
	3	8	6,991	2	3,009	10
	4	3	3,249	7	6,751	10
	5	1	1,013	9	8,987	10
	6	0	,266	10	9,734	10
	7	0	,039	10	9,961	10
	8	0	,005	10	9,995	10
	9	0	,001	10	9,999	10
	10	0	,000	10	10,000	10

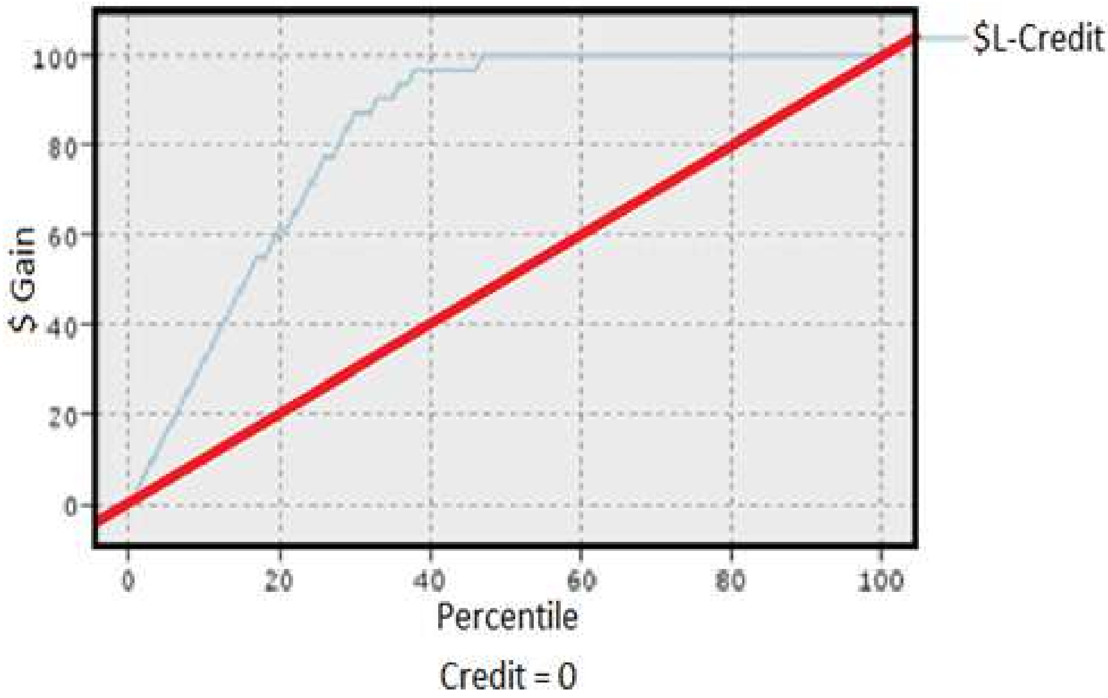
With the LR Model, it can be said that the situation of credit approving and not approving is sufficient model. The appropriateness of the LR Model is evaluated, the relation between the dependent variable and the independent variable is the most used statistics in the literature; McFaden R^2, Cox-Snell and Nagelkerke R^2 tests were used. The test results are shown in Table 9.

Table 9: Model R2 table.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	$37,492^a$	,578	,814

Negelkerke R^2 statistic table was found as ”0.814”. This shows that the variables used in the logistic model explain the model as ”81.4%”. In terms of modeling the variables, the LR model has a very good percentage. The Cox and Snell R2 values were also found to be ”0.578”. The explanatory ratios are quite high and it can be said that the established model is meaningful. The -2Log likelihood value is ”37,492”. The higher this value, the better it works.

Table 10: Performance graph of the LR analysis model.



The validity of the LR model established in the study was examined, it was observed that the model showed 92% success and was considered as a sustainable model.

Results for output field Credit

Comparing $L-Credit with Credit

Correct	92	92%
Wrong	8	8%
Total	100	

Performance Evaluation

0	1,062
1	0,298

5 Conclusions

The ”LR Model” was created to predict creditworthiness according to the identified fugitives. In the model, customer age, education, marital status, debt grade, credit card debt, other debts, cross product are the independent variables. These are statistically significant in terms of marital status, gender, cross product, also found that this affected creditworthiness positively. However, various variables such as debt-income ratio, credit card debt, and other debts are statistically significant and affect creditworthiness negatively. In addition, occupational, income and educational constraints were found to be meaningless.

In this research, it is aimed that the bank branch will infer the credits suitability from all customers in the customer portfolio by way of the personal information of the new and old customers to be included in the portfolio. With this model, the factors affecting the credit were evaluated. Factors influenced by credit approval; "Marital_Status (2)", "Gender(1)", "Cross_Product", "Year", "Debt_Income", 5 variables, were found to be significant. As a result of the study, it is found out that the credit approval rate of the customers is very high if the marital status of the customer is married out of the meaningful variables. The widow found that the approval rate of the customer's credit is low. It is found that the credit approval rate for male customers is high and female customers are low. The more Cross_Product the customer has, the more likely it affects the approval of the loan. The longer the customer is working with the bank, the more likely it is that credit approval is easier.

Also, the higher the "Debt_Income", the lower the approval of customers' credits will be reduced negatively. To become an example customer for this bank; work with the bank for many years, have a lot of cross products in the bank, debt is low, married and must be male. As a result of the study, the bank branch will benefit from the statistical model in which it is created, to evaluate according to the customer characteristics in its portfolio, and to give more credit to branch customers.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have contributed to all parts of the article. All authors read and approved the final manuscript.

References

[1] Atiya, Amir F. "Bankruptcy prediction for credit risk using neural networks: A survey and new results." *IEEE Transactions on neural networks* 12.4 (2001): 929-935.

[2] Maher, John J., and Tarun K. Sen. "Predicting bond ratings using neural networks: a comparison with logistic regression." *Intelligent systems in accounting, finance and management* 6.1 (1997): 59-72.

[3] Baesens, Bart, et al. "Benchmarking state-of-the-art classification algorithms for credit scoring." *Journal of the operational research society* 54.6 (2003): 627-635.

[4] Zekic-Susac, Marijana, Natasa Sarlija, and Mirta Bensic. "Small business credit scoring: a comparison of logistic regression, neural network, and decision tree models." *Information Technology Interfaces, 2004. 26th International Conference on*. IEEE, 2004.

[5] Bensic, Mirta, Natasa Sarlija, and Marijana Zekic-Susac. "Modelling small-business credit scoring by using logistic regression, neural networks and decision trees." *Intelligent Systems in Accounting, Finance and Management* 13.3 (2005): 133-150.

[6] Su-juan, P. A. N. G. "An application of logistic regression model in credit risk analysis." *Mathematics in Practice and Theory* 9 (2006): 020.

[7] Lee, Tian-Shyug, et al. "Mining the customer credit using classification and regression tree and multivariate adaptive regression splines." *Computational Statistics & Data Analysis*50.4 (2006): 1113-1130.

[8] Ata, H. Ali. "Banka Yabancilaşmasinin Türkiye'deki Yerli Ve Yabanci Bankalar Açisindan Karşilaştirilmasi." *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi* 23.4 (2009).

[9] Dong, Gang, Kin Keung Lai, and Jerome Yen. "Credit scorecard based on logistic regression with random coefficients." *Procedia Computer Science* 1.1 (2010): 2463-2468.

[10] Kong-lai, Zhu, and Li Jing-Jing. "Studies of discriminant analysis and logistic regression model application in credit risk for China's listed companies." *Management Science and Engineering* 4.4 (2010): 24.

[11] Nie, Guangli, et al. "Credit card churn forecasting by logistic regression and decision tree." *Expert Systems with Applications*38.12 (2011): 15273-15285.

[12] Budak, Hüseyin, and Semra Erpolat. "Kredi Riski Tahmininde Yapay Sinir Ağları ve Lojistik Regresyon Analizi Karşılaştırılması." *AJIT-e: Online Academic Journal of Information Technology* 3.9 (2012): 23-30.

[13] Kara, Mehmet, and Duran Kuru. "Satiş Geliştirme Faaliyetlerinin Satin Alma Kararina Etkisi: Hizmet Sektöründe Bir Uygulama." *Journal of Social Sciences/Sosyal Bilimler Dergisi* 6.12 (2013).

[14] Bekhet, Hussain Ali, and Shorouq Fathi Kamel Eletter. "Credit risk assessment model for Jordanian commercial banks: Neural scoring approach." *Review of Development Finance* 4.1 (2014): 20-28.

[15] Khemais, Zaghdoudi, Djebali Nesrine, and Mezni Mohamed. "Credit Scoring and Default Risk Prediction: A Comparative Study between Discriminant Analysis & Logistic Regression." *International Journal of Economics and Finance* 8.4 (2016): 39.

[16] Agresti, Alan. "An introduction to categorical data analysis, 1996." (1996): 34-39.

[17] Ayres, Gabriela, and Wei Wei. "Credit Scoring Model Applications: Testing Multinomial Targets." (2014).