

An Application of Multinomial Mis-Classification Cost Matrix For A P- To – P Lending Credit Score

Bakker, Daniel K.

University of Eastern Africa, Baraton

Ong'eta, Oyar Jackson

University of Eastern Africa, Baraton

ABSTRACT

An emerging new form of online credit for lending, different from traditional sources of finance, such as banks and building societies, where lenders provide loans to borrowers directly is termed as p^2 . Many of these credits are unsecured personal loans, thus credit score of loans is vital to regulate the default risk and improve profit for lenders and platforms. Standard two-fold classifiers may not be appropriate in this lending since there are multiple credit classes and misclassification costs vary largely across classes in the lending platforms. Cost Sensitive Classifiers have been studied extensively in this set of lending, but none of them have analyzed this issue from the perspective of multinomial classifications and measured the misclassification costs: $(C_{ab}^1)_{a>b}$ & $(C_{ab}^2)_{a<b}$, of different credit grades using actual losses and opportunity costs. The research intends to model credit score in p^2 lending as a cost-sensitive multinomial classification problem. A misclassification cost matrix is proposed for credit scoring with a set of equations and models to estimate the costs. A replication study using a publicly available data is conducted to evaluate the performance and validate the usefulness of the proposed misclassification cost matrix with the help of an R statistical package developed to aid the application of the model. The outcomes showed that the cost-sensitive multinomial classifiers can significantly decrease the total cost, which is vital for the p^2 survival and profitability.

Keywords: Misclassification Cost Matrix, Default Risk, P^2 , Two-fold Classifier, R Package.

INTRODUCTION

In the past years, online peer-to-peer (P^2) loaning, as a popular form of personal loan, has developed in credit market. It transfers traditional way of face-to-face personal loans through online services [1] It is cost effective, pervasive, convenient, efficient without the involvement of traditional financial institutions [11].

P^2 loaning, in comparison with the traditional banking system, has the following features.

- It facilitates transactions by linking borrowers and creditors directly. There are electronic loan application forms filled in by the borrowers, including amounts, terms, purposes, and personal information (such as age, job, address, and credit card among others). Platforms provide available financial conditions and credit histories of

borrowers to creditors, who will decide whether to grant a loan, the amount and at what interest rate. Platforms use various approaches to help lenders set interest rates. Some platforms carry out an auction at which a borrower set her/his maximum interest rate and creditors give their bids [9]. Another approach is to assign interest rates automatically using borrowers' credit grades, which are calculated based on borrowers' characteristics [6].

- P^2 loaning platforms charge service fees for transactions [13], instead of charging borrowers higher interest rates than the cost of the money as traditional financial institutions. P^2 loaning process benefits both borrowers and lenders. While borrowers can be granted money at lower costs than traditional financial institutions, creditors can make more money than putting their money in banks. This benefit comes with the risk of borrowers' defaulting on the loans because many P^2 loans are unsecured personal loans and most creditors may have little knowledge about credit risk management [21].

In order to regulate the default rates and risks, P^2 loaning platforms came up with classification models to evaluate credit risks of loans and borrowers and suggest appropriate interest rates for loan applications. The quality of these models is vital to the credit risk management and sustainability of P^2 lending platforms. Using experiences from financial institutions, P^2 loaning platforms adopt and develop classification algorithms to categorize borrowers into different credit grades based on their characteristics and credit history, and recognize potential borrowers who are likely to default [8,14].

Though it is a common practice in traditional credit rating to use standard cost-insensitive binary classification algorithms [15,16], such as logistic regression, neural networks, and decision trees [4], they are not appropriate in P^2 loaning due to the following reasons

- There are more than two classes of credit grades in P^2 lending and each credit grade implies a certain level of risk. This implies that multinomial classification should be considered in P^2 credit grading.
- P^2 credit data are imbalanced, meaning that number of samples in different credit grades varies. By illustration, the number of ideal borrowers in the best grade or high-risk borrowers in the worst grade is much smaller than the other grade groups.
- Misclassification costs are not uniform across classes in P^2 loaning. In general, the cost committing a type II error (classifying a loan with bad credit as a good one) is usually greater than type I error (classifying a good one as bad) [5].

In a multinomial credit-grading scenario, classifying a sample of grade C into grade A , for example, is more costly than classifying B into A . Therefore, standard cost-insensitive multinomial classification, in which all errors have the same cost, is not suitable for credit rating in P^2 loaning. Cost-sensitive multinomial classifiers would fit well for credit rating in P^2 loaning. These cost-sensitive classifiers were developed for imbalanced data classification [15]. Various cost-sensitive classifiers have been proposed for credit rating [2], with the aim of to minimize total costs measured by a misclassification cost matrix [10]. This is not only necessary but also important for cost-sensitive classification problems.

According to few studies in P^2 lending [21], few studies in P^2 loaning about cost-sensitive classifiers have been done, but none of them have analyzed this issue from the perspective of

multinomial classifications and measured misclassification costs of different credit grades with P^2 loaning.

Misclassification costs are losses of creditors' earnings due to misclassifying credit grades of loans. It equals to the difference between the return of a loan when it is correctly classified and the return of a loan when it is misclassified as other credit grade. This difference can be one of the following situations:

- i. If a loan is classified to a better credit grade with a lower interest rate, the risk to default of the loan is underestimated and the interest rate of the loan is set lower than it should be, which means that the interest maybe deficient to cover the risk that the lender bears. The creditor will lose potential returns that they could have gotten, including an unpaid risk that the borrower should pay for the higher-risk loan.
- ii. If a loan is classified to a worse credit grade with a higher interest rate, borrowers might be scared away or it may increase their chance to default, which causes opportunity costs and financial losses to lenders.

Objective of the Study

This paper proposes a multinomial cost matrix that measures misclassification costs of P^2 credit grading by considering real losses and opportunity costs associated with P^2 loaning. A set of equations and models were developed to calculate these misclassification costs. The parameters in the proposed equations and models are designed to calculate the cost matrix and support P^2 loaning platforms' operations. A case study using data from one of the approved Lending institutions within sub-Saharan region in Africa is conducted to demonstrate the performances of the proposed cost matrix using some well-known cost-sensitive classifiers. The results show that the proposed cost matrix can not only reveal the sources of losses caused by misclassifications, but also reduce the total costs for the P^2 platforms, which in comparison to the cost-insensitive classification algorithms, is better.

LITERATURE REVIEW

Most classifiers aim to maximize accuracy and minimize misclassifications. Various classification methods have been proposed for credit rating and risk management [17] Standard classifiers treat the costs of misclassifications the same, which should not true in real credit risk management. Many researches support the use of cost-sensitive classifiers in credit rating. Sahin et al. (2013) proposed a cost-sensitive decision tree approach with varying misclassification costs. It is successfully used in credit card fraud detection to decrease financial losses. Alejo, Garca, Marques, Sanchez, and Antonio-Velazquez (2013) improved the Multilayer Perceptron neural network using three misclassification cost functions and can be used to improve the prediction effectively in credit rating. [2] suggested example-dependent cost-sensitive methods and proposed logistic regression and decision trees for credit scoring.

Cost Matrices

Misclassification cost can be described by a cost matrix $C = (c_{ij})_{m \times m}$, where c_{ij} indicates the cost due to misclassifying an instance of class i as class j , and m is the number of classes. In credit rating, the measurement of misclassification costs in C is not only a basic component of cost-sensitive classification, but also vital for high quality credit rating. Real financial indicators, like profit-based or financial loss-related measures, are well aligned with the

objectives in credit rating [3,14,19] set the cost of a false negative to a loan's interest rate charged to the customer int_r , the cost of a false positive to the loss given default (LGD), and both the costs of true positive and true negative are set to zero

Table 1: Cost matrix proposed by Beling et al. (2005).

			Prediction	
			Positive	Negative
Actual	Positive		0	int_r
	Negative		LGD	0

As proposed by [12], a cost matrix (Table 2) for credit card fraud detection. It represents the costs of mis-classification by the administrative cost C_a , which is related to analyzing the transactions and contacting card holders. In the cases of false negative and true negative, the associated costs both equal to C_a because the card holder will have to be contacted. However, in the case of false positive, due to the fact that frauds are not detected, the cost is defined as a $100 C_a$.

Table 2: Cost matrix proposed by Hand et al. (2008).

			Prediction	
			Positive	Negative
Actual	Positive		0	C_a
	Negative		$100 * C_a$	C_a

[2] pointed out a limitation of the above cost matrices. Since losses of different frauds range from a few to a large amount, it is unrealistic to assume constant cost in false positive. In Bahnsen [2], a new cost matrix (Table 3) is proposed as a better representation of the actual costs, where the cost of false positive is replaced by the amount Amt_i of the transaction i .

Table 3: Cost matrix proposed by [2].

			Prediction	
			Positive	Negative
Actual	Positive		0	C_a
	Negative		Amt_i	C_a

He further proposed a cost matrix (Table 4) with example-dependent varying misclassification costs. For every borrower i , the costs of correct classifications are zero, and the cost of false positive is the losses if borrower i defaults which is proportional to his credit line Cl_i . The cost of false negative is $r_i + C^a$, where r_i is the profit that can be earned from a good borrower, and C^a is related to uncertainty of next alternative borrower.

Table 4: Example-dependent cost matrix proposed by [2].

			Prediction	
			Positive	Negative
Actual	Positive		0	$r_i + C^a$
	Negative		$Cl_i * LGD$	0

[21] proposed a cost-sensitive boosted tree loan evaluation model to discriminate potential default borrowers in P2P lending. The model considers the imbalanced misclassification cost (Table 5), based on the assumption that the cost of misclassifying a default borrower is larger than that of misclassifying a good one, i.e., $C(0, 1) > C(1, 0)$. It adopts a cost matrix with Default and non-default classes.

Table 5: Cost matrix in Xia et al. (2017).

			Prediction	
			Non-default	Default
Actual	Non-default		0	$C(1,0)$
	Default		$C(0,1)$	0

METHODOLOGY

Modeling in P² Lending

P² lending platform provides services to creditors and borrowers to facilitate their transactions, and charges a proportion of profits earned by successful repaid loans as an essential part of services fee. In such a case, the platforms and lenders share profits and risks. In this paper, loss given default (LGD) as per [18] is used to measure the proportion of a lender's loss when a borrower defaults on a loan. Although LGD is different for different loans, it is usually treated as a constant number for all loans to simplify models [2,3]. For platforms, default loans impact their profits directly by causing the loss of service fees and indirectly by damaging their reputations in quality of services on credit rating and setting interest rates. Therefore, accurate credit rating is the foundation of healthy and profitable operations of P² platforms. Consequently, an important requirement of credit rating in P² loaning is to measure the profitability of a loan [21], and the amount of loss in profit for creditors should be considered in misclassification cost matrix.

P² platforms usually classify borrowers into multiple credit grades and determine hierarchical interest rates for them. Table 6 shows the constructed structure of credit grades to measure the profits and losses in P² loaning, based on the assumption that different credit grade i is associated with a certain probability of default (PD) PD_i . Therefore, it is assigned an interest rate I_i to reflect the corresponding credit risk. Credit rating classifies loans into these grades, where 1(A) refers to the best credit, and n refers to the worst credit risk. Generally, the worse the credit grade, the higher probability of default and the higher interest rate. That is, for any $i, j = 1, 2, 3, \dots, n$, $i > j$, so that $PD_i > PD_j$ and $I_i > I_j$.

When a lender lends a loan of credit grade i with PD_i and I_i , the lenders' expected return is:

$$\begin{aligned} ER_i &= (1 - PD_i)(1 + I_i) + PD_i(1 - LGD) \\ &= 1 + I_i - PD_i(I + LGD) \end{aligned} \quad (1)$$

Table 6: Proposed Structure of Credit Grades in P² loaning.

Credit Grade (CG)	Probability of Default (PD)	Interest Rate (I)
1(A)	PD_1	I_1
2(B)	PD_2	I_2

3(C)	PD_3	I_3
4(D)	PD_4	I_4
...	...	...
n	PD_n	I_n

Misclassification Cost Matrix

Table 7 shows a proposed misclassification cost matrix C , for credit rating in P^2 loaning where the cost of correct classification is zero, i.e., $c_{ii} = 0$, based on Table 6. This Misclassification cost

matrix C can be decomposed into two blocks C^1 and C^2 , that is: $C = \begin{pmatrix} 0 & C^2 \\ C^1 & 0 \end{pmatrix}$. C^1 corresponds

to the loans whose predicted credit grades (j) are better than their actual credit grades (i), i.e., $i > j$, and C^2 is the upper triangular submatrix corresponding to the loans whose actual credit grades (i) are better than their predicted credit grades (j), i.e., $i < j$.

Table 7: Misclassification cost matrix.

		Prediction				
		1(A)	2(B)	3(C)	...	n
Actual	1(A)	0	c_{12}	c_{13}	...	c_{1n}
	2(B)	c_{21}	0	c_{23}	...	c_{2n}
	3(C)	c_{31}	c_{32}	0	...	c_{3n}
	...				0	...
	n	c_{n1}	c_{n2}	c_{n3}	...	0

Lower Triangular Submatrix:

In the lower triangular matrix $C^1 = (C_{ij}^1)_{i>j}$, C_{ij}^1 is the cost of misclassifying a loan in grade i as grade j . As a result of misclassification, $PD_i > PD_j$; $I_i > I_j$. This means a borrower pays a lower interest rate. Lower interest rate may affect borrowers' potential defaults to a certain extent, which has been supported by theoretical and empirical evidences [7,19]. This kind of changes can be described by revising the PD in a linear function related with the difference of interest rates:

$$PD'(j|i)_{i>j} = PD_i + \beta(I_i - I_j) \quad (2)$$

Where the parameter β (< 0), is a revised coefficient of PD , indicating how the PD of a borrower changes with different interest rates. $PD'(j|i)_{i>j}$ is the actual probability of default, after misclassifying a loan in grade i to grade j . The misclassification leads to a change of equation (1) as:

$$ER'(j|i)_{i>j} = 1 + I_j - PD'(j|i)(I_j + LGD) \quad (3)$$

The misclassification also leads to a loss in return cost formulated as:

$$C_{ij}^1 = ER_i - ER'(j|i) \\ = (I_i - I_j)[1 - PD_i + \beta(I_i + LGD)] \quad (4)$$

Upper Triangular Submatrix:

In the upper triangular matrix $C^2 = (C_{ij}^2)_{i < j}$, $PD_i < PD_j; I_i < I_j$, which means a borrower pays a higher interest rate. This may lead to a withdrawal of loan application with a certain probability because the borrower can't endure the high interest rate. This probability of application withdrawal is related to the difference of interest rates $I_i - I_j$. To simplify the model, a linear function is used to calculate the probability that a borrower of credit grade i will give up application when misclassified as credit grade j (worse than grade i):

$$\Pr_{gu}(j|i)_{i < j} = \alpha(I_i - I_j) \quad (5)$$

where the parameter α is borrowers' attrition rate [23], which indicates how the probability $\Pr_{gu}(j|i)_{i < j}$ varies with the difference of interest rates $I_i - I_j$.

On the other hand, the probability of a borrower accepting interest rate I_j caused by misclassification is $1 - \Pr_{gu}(j|i)$. In this case, the PD of the borrower increases with the higher interest rate. Equation (6) is used to calculate the actual PD $PD'(j|i)_{i < j}$ when $i < j$.

$$PD'(j|i)_{i < j} = PD_i + \beta(I_i - I_j) \quad (6)$$

For a lender who accepts the misclassified loans when $i < j$, the expected return is:

$$ER'(j|i)_{i < j} = 1 + I_j - PD'(j|i)(I_j + LGD) \quad (7)$$

Thus, the misclassification cost C_{ij}^2 can be measured as:

$$C_{ij}^2 = ER_i - (1 - \Pr_{gu}(j|i))ER'(j|i) \quad (8)$$

Equation 8 can be rewritten as:

$$C_{ij}^2 = (1 - \Pr_{gu}(j|i))(ER_i - ER'(j|i)) + \Pr_{gu}(j|i)ER_i \quad (9)$$

In equation (9), the first term on the right-hand side indicates the loss in profit from transactions, and the second term indicates the opportunity cost caused by application withdrawals due to misclassifications. Based on a structure of credit grades in P^2 lending, we designed a misclassification cost matrix to measure real business losses, which are calculated using Equations (4) and (9).

Parameters Analysis

The revised coefficient for the probability of default, β , indicates how the PD of a borrower changes with different interest rates she/he bears. When $i > j$, even if a borrower of credit grade i lowers her/his PD from PD_i to $PD'(j|i)$, $PD'(j|i)$ is still larger than PD_j due to the lower interest rate I_j . This implies that $PD_j < PD'(j|i) < PD_i$. In reference to equation 2, we can write this as:

$$PD_j < PD_i + \beta(I_i - I_j) < PD_i, \text{ and since } I_i - I_j > 0, \text{ and for } i > j$$

$$\frac{PD_i - PD_j}{I_i - I_j} < \beta < 0 \quad (11)$$

According to equation 4, in view of cost matrix, provided that $i > j$ for $i, j = 1, 2, 3, \dots, n$

$$(I_i - I_j)[1 - PD_i + \beta(LGD + I_j)] > 0$$

$$\beta(LGD + I_j) + 1 - PD_i > 0$$

$$\beta > \frac{PD_i - 1}{LGD + I_j} \quad (12)$$

Moreover, for any $i, j, k = 1, 2, 3, \dots, n$, provided that $i > j > k$, the cost due to misclassifying a loan of credit grade i into k should be larger than misclassifying i as j .

If $C_{ik}^1 > C_{ij}^1$, then

$$I_i - I_k[1 - PD_i + \beta(LGD + I_k)] > I_i - I_j[1 - PD_i + \beta(LGD + I_j)]$$

$$= (I_j - I_k)(1 - PD_i) > (I_i - I_j)\beta(LGD + I_j) - (I_i - I_k)\beta(LGD + I_k)$$

If we set $(I_i - I_j)(LGD + I_j) > (I_i - I_k)(LGD + I_k)$, then

$$\frac{(I_j - I_k)(1 - PD_i)}{(I_i - I_j)(LGD + I_j) - (I_i - I_k)(LGD + I_k)} > \beta$$

If we set $(I_i - I_j)(LGD + I_j) < (I_i - I_k)(LGD + I_k)$, then

$$\frac{(I_j - I_k)(1 - PD_i)}{(I_i - I_j)(LGD + I_j) - (I_i - I_k)(LGD + I_k)} < \beta \quad (13)$$

If $C_{ik}^1 > C_{jk}^1$, then

$$\frac{(I_j - I_k)(1 - PD_i) - (1 - PD_i)(I_i - I_k)}{(I_i - I_j)(LGD + I_k)} < \beta \quad (14)$$

The borrower's churn rate, α , indicates how the probability $\Pr_{gu}(j|i)$ which is within the interval $(0,1)$ varies with the interest rates difference $I_i - I_j$ for $i, j=1,2,3,\dots,n$ and $i < j$. based on the interval,

$$\frac{1}{I_i - I_j} < \alpha < 0 \quad (15)$$

$$C_{ij}^2 > 0 \quad (16)$$

The cost of misclassifying a bad loan into a better one is larger than the other way around. That is

$$C_{ji}^1 > C_{ij}^2 \quad (17)$$

And for $i < j < k$

$$C_{ik}^2 > C_{ij}^2 > C_{jk}^2 \quad (18)$$

DISCUSSION OF THE RESULTS

Implications of the Parameters

The interest rates in P^2 loaning are set based on the business environment, which can be measured using parameters, such as α and β . The range of parameter β is solved as $-0.11 < \beta < 0$ to satisfy equations (11)– (14). We then set $\beta = -0.56$, that is, the median of the range, in order to get the proposed model. Thus, by applying equation 2, we get the results shown in Table 11. Bold numbers indicate the inherent PD s for loans of credit grades. In fact, fluctuations of interest rates caused by misclassification can affect borrowers' default. The experiment proved that inherent credit ratings have more effects on the probability of default than the fluctuations of interest rates. This makes sense in P^2 lending, where loans are normally small in size and repayment of unsecured microloans is more dependent on the willingness to pay than the ability to pay.

The range of parameter is solved as $-1.30 < \alpha < 0.84$ to satisfy equations (15) to (19). We set α as the median of the range, this is $\alpha = -1.07$ to satisfy equations in 15 to 19. It indicates that a borrower's probability of terminating a loan application increases about 1.1% if the interest rate rises by 1%. Should there be a big difference between the values of parameters calculated and the results obtained from the real practice, the interest rates on P^2 lending platforms should be adjusted, because the given PD 's are generally stable.

Table 11: Revised PD on lending club.

		Prediction						
		A	B	C	D	E	F	G
Actual	A	0.062	0.065	0.067	0.068	0.069	0.070	0.071
	B	0.113	0.115	0.117	0.118	0.119	0.120	0.121
	C	0.152	0.154	0.156	0.157	0.158	0.159	0.161
	D	0.179	0.182	0.183	0.185	0.186	0.186	0.187
	E	0.201	0.203	0.204	0.206	0.207	0.207	0.208
	F	0.243	0.246	0.247	0.248	0.249	0.250	0.251
	G	0.331	0.333	0.335	0.336	0.337	0.338	0.339

Cost Matrix

We calculated the misclassification cost matrix C in (Table 12) according to Equations (4) and (9). Elements in the cost matrix measure the loss of creditors' expected return if a loan is misclassified. For example, in the cost matrix C , $c_{21} = 0.033$ means that a creditor who provides a loan with a certain amount (amt) will loss $0.033 \cdot amt$ in the expected return when a loan of credit grade B is misclassified as grade A . unfortunately, the revenue of the platform reduces because of the misclassification of credit grades, where the costs should be minimized in cost-sensitive credit rating.

The figure 0.086 happens when class G was misclassified as class A . This misclassification assigns the lowest interest rate to the highest credit risk group. And the smallest cost is 0.0055, which appears when class E is misclassified into F .

Table 12: Misclassification cost matrix on lending club.

		Prediction						
		A	B	C	D	E	F	G
Actual	A	0	0.009	0.017	0.024	0.030	0.037	0.044
	B	0.033	0	0.007	0.014	0.020	0.026	0.034
	C	0.053	0.021	0	0.007	0.013	0.018	0.026
	D	0.068	0.038	0.017	0	0.006	0.011	0.018
	E	0.079	0.049	0.029	0.012	0	0.006	0.012
	F	0.085	0.057	0.039	0.022	0.011	0	0.007
	G	0.086	0.061	0.044	0.030	0.020	0.011	0

To facilitate observations of the cost matrices under different parameters, cosine similarity is adopted to compare them with the cost matrix $C = (c_{ij})$ shown in Table 12, where parameters are taken as median (α, β) . Given any $\alpha = (-1.29, -0.84)$ and $\beta = (-0.11, 0)$, the corresponding cost matrix, $M = (m_{ij})$ can be obtained. Then the cosine similarity between C and M is calculated as:

$$\text{Cosine Similarity } (C, M) = \frac{\sum c_{ij} m_{ij}}{\sqrt{\sum c_{ij}^2 \cdot \sum m_{ij}^2}} \quad (20)$$

The cost matrix C in Table 12 is considered as benchmark in calculation of cosine similarity.

Table 13 lists the cosine similarities on these discrete values of the parameters α and β taken at 5 equidistant points within their ranges. When both parameters α and β are taken as the maximum in their ranges, the fact that similarity equals 0.945 still guaranteed a small difference between the corresponding cost matrix in Table 13 and benchmark in Table 12.

It is reasonable to set the parameters as the median value of the ranges since the sensitivity analysis for the parameters α and β shows that the proposed method used to calculate cost matrix is robust and stable in the presence of uncertainty of parameters.

Table 13: Cosine similarities of cost matrices.

Alpha \ Beta	-0.112	-0.084	-0.056	-0.028	0.000
-1.292	0.962	0.965	0.968	0.971	0.973
-1.179	0.987	0.989	0.991	0.992	0.993
-1.066	1.000	1.000	1.000	1.000	1.000
-0.954	0.990	0.989	0.988	0.987	0.986
-0.841	0.948	0.947	0.947	0.946	0.945

Cost-Sensitive Credit Rating

In order to compare the performance, we selected four classifiers [14] and the corresponding Meta-cost-sensitive classifiers(C-S). The four classifiers are LDA, Random Forest (RF), Logistic Regression (LR), and Support Vector Machines (SVM). The performances of classifiers were evaluated using accuracy, total cost, average cost, and cost saving rate. They were computed using the confusion matrix N in classification result and the cost matrix.

The classification results were summarized in Figure 1.

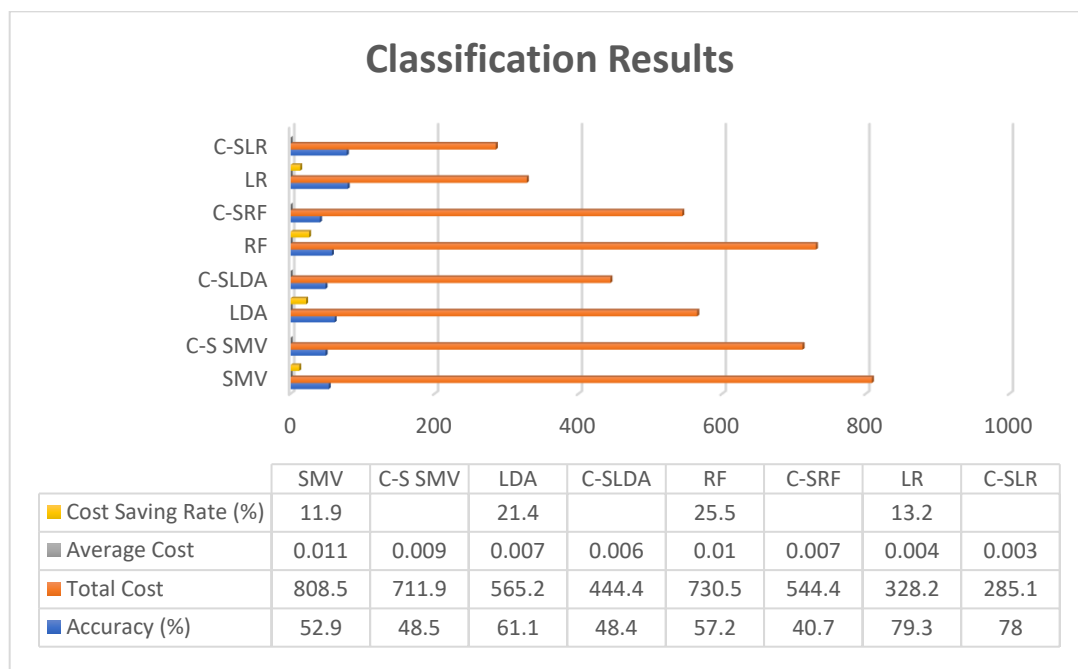


Figure 1: Classification results

Since traditional standard classifiers focus on optimizing the accuracy and ignoring the relationship between classes and different misclassifications, they have higher accuracies than their cost-sensitive counterparts, and the cost-sensitive classifiers reduced the total costs dramatically, which is a major concern of P^2 loaning platforms and creditors.

In order to achieve the highest accuracy and the lowest total cost simultaneously is really difficult, if not impossible. Cost-sensitive classification normally sacrifices accuracy for lower total cost [20]. The objective of cost-insensitive classifiers is to maximize the total accuracy. But none of them can reach 100% accuracy in multinomial classification. Although misclassified instances may cause different costs according to the cost matrix, cost-insensitive classifiers treat them the same. This is the reason that cost-insensitive classifiers have higher accuracy and higher total misclassification cost. Cost-sensitive classifiers, on the other hand, try to minimize total costs caused by misclassification, and some misclassification errors with low costs are compromised to achieve this goal when training the classifiers. Thus, cost-sensitive classifiers reduce total cost of classification, but have lower accuracy than cost-insensitive classifiers.

As shown in Figure 1, the classifiers which achieved the highest accuracy and the lowest total cost are LR and cost-sensitive LR, respectively. While the accuracy of cost-sensitive LR is 1.3% lower than LR, cost-sensitive LR reduced 13.2% total cost, compared to LR.

Since the proposed cost matrix measured the real losses of creditors, the total cost is more practical than the accuracy. In order to connect the classification results and business reality in P^2 loaning, the average cost indicates how much money is lost in a loan, on average, due to potential misclassification of credit grades when a classification algorithm is used. For example, if standard LR is used to grade a loan, potential misclassification cost will cause 0.4% less returns for creditors, comparing to 100% accurate classification. And this cost dropped to 0.3% if cost-sensitive LR is used.

Applying cost-sensitive classifiers for credit rating is able to reduce losses effectively in P^2 lending, which is measured by the cost matrix proposed in this paper. It also shows the need to evaluate and develop more credit rating approaches with respect to minimizing financial losses for P^2 lending platforms.

CONCLUSION AND RECOMMENDATION

Cost-sensitive classifiers try to minimize total costs measured by misclassification cost matrixes. How to measure misclassification costs of different credit grades in P^2 loaning is a useful but understudied issue. We first proposed a misclassification cost matrix for P^2 credit grading that takes into account real losses and opportunity costs associated with P^2 loaning. A set of equations were developed to calculate the costs in the misclassification cost matrix. Then parameters were analyzed using the proposed equations and models from the perspective of both cost-sensitive classification and business operations.

The results showed that standard classifiers have higher accuracy than cost-sensitive counterparts, but the cost-sensitive classifiers significantly reduced the total costs, which is essential for the survival and profitability of P^2 loaning platforms.

The parameter *LGD* can be included in the future research together with the administrative costs in the total cost. Since the administrative cost in P^2 loaning, comparing with traditional bank loans, is insignificant, the reduced total cost calculated in the experiment may still be used as a useful reference for P2P loaning platforms. Future research can also investigate how changes in interest rates affect human behaviors on default and giving up a loan.

References

1. Bachmann, A., Becker, A., Buerckner, D., Hilker, M. K. M., Lehmann, M., & Tiburtius, P. (2011). Online peer-to-peer lending-a literature review. *Journal of Internet Banking and Commerce*, 16(2), 1–18.
2. Bahnsen, A. C., Aouada, D., & Ottersten, B. (2015). Example-dependent cost-sensitive decision trees. *Expert Systems with Applications*, 42(19), 6609–6619. doi:10. 1016/j.eswa.2015.04.042
3. Beling, P., Covaliu, Z., & Oliver, R. M. (2005). Optimal scoring cutoff policies and efficient frontiers. *Journal of the Operational Research Society*, 56(9), 1016–1029. doi: 10.1057/palgrave.jors.2602021
4. Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance*, 72, 218–239. doi: 10.1016/j.jbankfin.2016.07.015
5. Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: A recent review. *Artificial Intelligence Review*, 45(1), 1–23. doi:10.1007/s10462-015-9434-x
6. Collier, B., & Hampshire, R. (2010). Sending mixed signals: Multilevel reputation effects in peer-to-peer lending markets. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (pp. 197–206). ACM. doi:10.1145/1718918.1718955
7. Edelberg, W. (2006). Risk-based pricing of interest rates for consumer loans. *Journal of Monetary Economics*, 53(8), 2283–2298. doi: 10.1016/j.jmoneco.2005.09.001
8. Florez-Lopez, R., & Ramon-Jeronimo, J. M. (2014). Modelling credit risk with scarce default data: On the suitability of cooperative bootstrapped strategies for small low-default portfolios. *Journal of the Operational Research Society*, 65(3), 416–434. doi:10.1057/jors.2013.119
9. Galloway, I. (2009). Peer-to-peer lending and community development finance. *Community Investments*, 21(3), 19–23.
10. Guan, D., Yuan, W., Ma, T., Khattak, A. M., & Chow, F. (2017). Cost-sensitive elimination of mislabeled training data. *Information Sciences*, 402, 170–181. doi:10. 1016/j.ins.2017.03.034
11. Guo, Y., Zhou, W., Luo, C., Liu, C., & Xiong, H. (2016). Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research*, 249(2), 417–426. doi:10.1016/j.ejor.2015.05.050
12. Hand, D. J., Whitrow, C., Adams, N. M., Juszczak, P., & Weston, D. (2008). Performance criteria for plastic card fraud detection tools. *Journal of the Operational Research Society*, 59(7), 956–962. doi: 10.1057/palgrave.jors.2602418
13. Klafft, M. (2008). Peer to peer lending: Auctioning micro-credits over the internet. In *Proceedings of the International Conference on Information Systems, Technology and Management*. Dubai: IMT.
14. Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124–136. doi: 10.1016/j.ejor.2015.05.030

15. Li, T., Kou, G., Peng, Y., & Shi, Y. (2017). Classifying with adaptive hyper-spheres: An incremental classifier based on competitive learning. In *IEEE Transactions on Systems, Man, and Cybernetics: Systems* (pp. 1–12). IEEE
16. Morente-Molinera, J. A., Mezei, J., Carlsson, C., & Herrera-Viedma, E. (2017). Improving supervised learning classification methods using multi-granular linguistic modelling and fuzzy entropy. *IEEE Transactions on Fuzzy Systems*, 25(5), 1078–1089.
17. Santana, P. J., Lanzarini, L., & Bariviera, A. F. (2018). Fuzzy Credit Risk Scoring Rules using FRvarPSO. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 26(Suppl. 1), 39–57. doi:10.1142/S0218488518400032
18. Schuermann, T. (2004). What do we know about loss given default?
19. Serrano-Cinca, C., Gutierrez-Nieto, B., & Lopez-Palacios, L. (2015). Determinants of default in P2P lending. *PloSOne*, 10(10), e0139427 doi: 10.1371/journal.pone.0139427
20. Wang, H., Kou, G., & Peng, Y. (2018). Cost-sensitive classifiers in credit rating: A comparative study on P² lending. In *2018 7th International Conference on Computers Communications and Control (ICCCC)* (pp. 210–213). IEEE. doi:10.1109/ICCCC.2018.8390460
21. Xia, Y., Liu, C., & Liu, N. (2017). Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending. *Electronic Commerce Research and Applications*, 24, 30–49. doi:10.1016/j.elerap.2017.06.004
22. Xu, J., Chen, D., & Chau, M. (2016). Identifying features for detecting fraudulent loan requests on P2P platforms. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI)* (pp. 79–84). IEEE.
23. Zhu, B., Baesens, B., Backiel, A., & Vanden Broucke, S. K. (2018). Benchmarking sampling techniques for imbalance learning in churn prediction. *Journal of the Operational Research Society*, 69(1), 49–65. doi:10.1057/s41274-016-0176-1