Reinforcement Learning on the Credit Risk-Based Pricing

Tri Handhika
Centre for Computational Mathematics Studies
Gunadarma University
Depok, Indonesia
trihandika@staff.gunadarma.ac.id

Ahmad Sabri
Department of Informatics
Gunadarma University
Depok, Indonesia
sabri@staff.gunadarma.ac.id

Murni
Centre for Computational Mathematics Studies
Gunadarma University
Depok, Indonesia
murnipskm@staff.gunadarma.ac.id

***Abstract*—Credit scoring is the main process of credit transactions in assessing the credit risk of credit applicants. Unfortunately, in practice, its implementation only stops to the credit approvals. In this research, we utilize credit scores to generate the customized credit prices. We believe that each person has their own credit risk so that they will get different credit prices depend on their individual credit risk. This credit risk-based pricing is optimized by reinforcement learning approaches to represent the dynamic solution related to the updated credit historical data. There are several variables considered in the profit optimization model such as credit scores, tenor, credit prices (or rate for credit applicants), and plafond. We implement this solution to the random generated credit data.**

***Index Terms*—reinforcement learning, customized credit price, risk-based pricing, simulation, optimization**

I. Introduction

Risk analysis is an important part of the credit approval process as shown in Figure 1. We can see that a standardized credit rating is needed as a consideration for decision makers in approving credit application. It is obtained through the credit risk assessment model known as credit scoring [2]. Currently, the credit score is not only used for credit approval but also to determine credit prices known as credit risk-based pricing [3]. It can help credit companies in determining their optimal credit allocation [4].

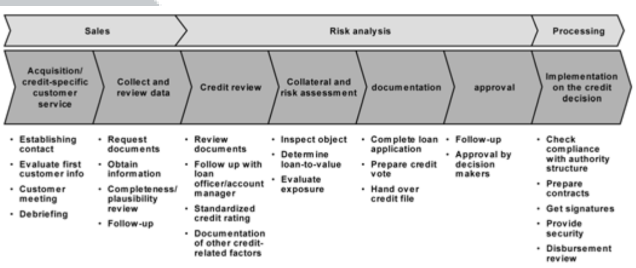

Fig. 1. Credit approval process [1].

The idea of credit risk-based pricing comes from the fact that every customer is unique. The feasibility of the customers to receive a loan is represented by their credit scores, assigned by the company based on their risk profile and previous credit performance (if any). Based on that, the loan rate is customized for each customer. This is agreed by Walke et al. to be implemented in pricing credit union loans in the United States [5]. Edelberg shows that lenders increasingly used credit risk-based pricing in consumer loan market during the mid-1990s [6]. Deng and Gabriel also proposed to apply this credit risk-based pricing to the mortgage credit [7, 8]. On the contrary, if the company gives the same average rates to every customer, then the low-risk customers will seek for other company that would give them better rates, while the relatively high-risk customers stay and accept the price. This has an impact on disparity in consumer credit history scores [9] such that will increase average losses and reduce the profitability for the company.

Theoretically, low-risk customers pay fewer rates than those of the higher-risk. The question is, given the price elasticity, how much should the loan rate be assigned to each customers segment to maximize the company's profitability? Through credit scoring and credit pricing dataset consisting of tenor, credit prices (or rate for credit applicants), and plafond, we propose a reinforcement learning method to find the solution of this problem.

To determine optimal prices from each existing segment the net interest income can be combined with price elasticity as introduced in [10] when model the price optimization. The optimal profitability can be obtained by optimizing profit for each segment. This process is known as profit-based pricing. It combines not only the elasticity of price and cost, but also the risks that arise when maximize the profit [11].

There are many cases showing that credit companies can increase their profitability by conducting an analysis of credit pricing. One of the is the use of the Target Pricing System (TPS) in increasing profit of a UPS package shipper to more than $100 million per year as reported by Boyd et al. [12]. Increasing profitability also occurred in one of the subprime auto lenders, i.e. AmeriCredit, which implement the credit pricing optimization system [13]. In three months, AmeriCredit has increased their profit about $4 million.

In this paper we use reinforcement learning to accommodate the change of the price elasticity in current situation, or

This document is the results of the research project funded by Direktorat Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi / Badan Riset dan Inovasi Nasional (Grant No. 234/SP2H/LT/DRPM/2021).

978-1-6654-5866-5/21/$31.00 ©2021 IEEE
DOI 10.1109/ICCMST54943.2021.00056

specifically, the applicants' acceptance rate for a given loan rate. This approach to credit risk management is in line with current technological advances in the era of big data [14]. The massively increasing data requires the proper automation of the learning system [15]. For this purpose, we adapt the Kalman filter approach as provided in [16] to the model given in [10].

II. The Model and the Reinforcement Learning Scheme

We categorize the borrower based on plafond, tenor, credit scores. The present value of net interest income (PVNII) for amount borrowed P, tenor n, loan rate r and cost of capital r_c is approximated by

$$\text{PVNII} = Pn\left(r - r_c\right) - PD \times LGD, \qquad (1)$$

where PD and LGD stand for probability of default and loss given default, respectively [10].

We do not require the probability that the borrower will default at time j, $1 \leq j \leq n$, $(p_1, p_2, \cdots, p_n)$, in approximating equation (1) because PVNII is linear in rate. In other words, it means that we only need PD and LGD to approximate PVNII. Based on past observations, they are more observable and more stable than the full default probability vector. Therefore, we prefer to approximate PVNII in equation (1).

The total profit (π) is then optimized by (see [10]):

$$\max_{r} \pi(r) = \Sigma_i D_i \bar{F}_i\left(r_i\right)\left(\text{PVNII}\left(P_i, r_i, n_i\right) + v_i\right), \qquad (2)$$

$$\text{subject to } r_i \geq 0,$$

where:

- $N \geq 1$ is the number of pricing segments;
- $r_i = r_1, r_2, \cdots, r_N$ is the vector of rates offered to each pricing segment;
- $D_i > 0$ is the total demand (in number of loans) in pricing segment i;
- $P_i > 0$ is the average loan size in pricing segment i;
- $n_i > 1$ is the available term in pricing segment i;
- $\text{PVNII}\left(P_i, r_i, n_i\right)$ is the present value of net interest income;
- v_i is the present value of expected non-interest items (in segment i);
- $\bar{F}_i\left(r_i\right)$ is the acceptance rate.

In this paper, we used nonlinear programming to solve the optimization problem. We know that $\text{PVNII}\left(P_i, r_i, n_i\right)$ has a unique solution because it is log concave in r [10]. Noted that log concavity is weaker than concavity condition.

For different segments, the borrower price also has different sensitivity. It is determined not only by the loan size, but also the borrowers' risk level. The larger loans will be more price-sensitive than the smaller loans, while the higher risk of borrowers tend to be less price sensitive than the lower risk of borrowers. The Kalman filter is then applied to incrementally improve the estimation for the acceptance rate $\hat{\bar{F}}_i\left(r_i\right)$ for a given $r_{i,t}$ along the time step. The procedure is derived from that proposed by Carvalho [16]. For this purpose, the model needs the dataset that consists of plafond, tenor, grade of credit scores, and segment number.

The constraint in equation (2) suggests that all segments are independent to each other. It follows that the maximum profit is obtained by maximizing each segment independently. Specifically, the optimal profit for i-th segment at time t is given by

$$\max_{r_{i,t}} \pi_{i,t}(r_{i,t}) = D_{i,t}\bar{F}_{i,t}\left(r_{i,t}\right)\left(\text{PVNII}\left(P_i, r_{i,t}, n_i\right) + v_{i,t}\right), \qquad (3)$$

$$\text{subject to } r_{i,t} \geq 0.$$

The maximum profit is obtained when $r_{i,t} = r^*_{i,t}$. The total profitability formula for all segments at time t is easily constructed from equation (3), and we leave it to the reader. The observed acceptance rate $\bar{F}_{i,t}$ for i-th segment at time t is obtained by

$$\bar{F}_{i,t} = \frac{1}{1 + \exp\left[-\left(\alpha_i + \beta_i r_{i,t} + \epsilon_{i,t}\right)\right]}. \qquad (4)$$

Taking the logarithm of equation (4) we obtain a loglinear relation $y_{i,t} = -\ln \frac{1}{\bar{F}_{i,t}} - 1 = \alpha_i + \beta_i r_{i,t} + \epsilon_{i,t}$. The prediction for the acceptance rate for i-th segment at time t is given by

$$\hat{\bar{F}}_{i,t} = \frac{1}{1 + \exp\left[-\left(\hat{\alpha}_{i,t} + \hat{\beta}_{i,t} r_{i,t}\right)\right]}. \qquad (5)$$

Analogously, we have

$$\hat{y}_{i,t} = \hat{\alpha}_{i,t-1} + \hat{\beta}_{i,t-1} r_{i,t}. \qquad (6)$$

The estimated parameters $\hat{\alpha}_{i,t}$ and $\hat{\beta}_{i,t}$ in equation (5) or (6) are obtained from linear regression on prior data. To estimate α_i and β_i, we first consider a prior data having T records. Let assume $\theta_i = \left(\begin{smallmatrix}\alpha_i \\ \beta_i\end{smallmatrix}\right) \sim N\left(\hat{\theta}_{i,t}, \sigma_i^2 R_{i,t}\right)$ for $t = T$, where $\hat{\theta}_{i,t} = \left(\begin{smallmatrix}\hat{\alpha}_{i,t} \\ \hat{\beta}_{i,t}\end{smallmatrix}\right)$ and $\sigma_i^2 R_{i,t}$ is a covariance matrix (more precisely, σ_i^2 is an error variance and $R_{i,t} = \left(\begin{smallmatrix}1 & 1 & \cdots & 1 \\ r_{i,1} & r_{i,2} & \cdots & r_{i,t}\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & r_{i,1} \\ 1 & r_{i,2} \\ \vdots & \vdots \\ 1 & r_{i,t}\end{smallmatrix}\right)$).

In the next time step (algorithmically, $t = t + 1$), we find an optimal $r_{i,t} = r^*_{i,t}$; then define $z_{i,t} = \left(\begin{smallmatrix}1 \\ r_{i,t}\end{smallmatrix}\right)$, and apply the Kalman filter to update the current parameters by the following equations:

$$\hat{\theta}_{i,t} = \hat{\theta}_{i,t-1} + R_{i,t} z'_{i,t} S^{-1}_{i,t}\left(y_{i,t} - \hat{y}_{i,t}\right), \qquad (7)$$

$$R_{i,t} = R_{i,t-1} - R_{i,t-1} z'_{i,t} S^{-1}_{i,t} z_{i,t} R_{i,t-1}, \qquad (8)$$

$$S_{i,t} = z_{i,t} R_{i,t-1} z'_{i,t} + H_t, \qquad (9)$$

where $H_t = 1$. It is worth to mention that $\sigma_i^2\left(R_{i,t}\right)_{2,2}$ the variance of the estimated $\hat{\beta}_{i,t}$. For computational purposes, the reinforcement learning scheme on the credit risk-based pricing is represented in equations flowchart as shown in Figure 2.

TABLE I

DATASET USED FOR SETTING INITIAL PARAMETERS (R: RATE, D: DEMAND, F: ACCEPTANCE RATE, PD: PROBABILITY OF DEFAULT, LGD: LOSS GIVEN DEFAULT).

Grade of Credit Score	A					B					⋯	C				
Segment (i)	1					2					⋯	27				
Plafond (P_i) in Million Rupiahs	50					50					⋯	150				
Tenor (n_i) in years	1					1					⋯	3				
Index	r (%)	D	F (%)	PD (%)	LGD (mil. Rp.)	r (%)	D	F (%)	PD (%)	LGD (mil. Rp.)	⋯	r (%)	D	F (%)	PD (%)	LGD (mil. Rp.)
-400	20	59	50	6	2.50	22	37	79	9	4.50	⋯	48	82	52	49	73.50
-399	22	98	45	6	2.50	23	93	77	9	4.50	⋯	50	97	50	49	73.50
-398	18	54	56	6	2.50	18	33	81	9	4.50	⋯	48	14	52	49	73.50
-397	22	14	44	6	2.50	23	84	78	9	4.50	⋯	47	55	53	49	73.50
-396	20	7	51	6	2.50	21	63	79	9	4.50	⋯	49	22	51	49	73.50
-395	19	20	53	6	2.50	20	65	79	9	4.50	⋯	42	83	58	49	73.50
-394	21	47	46	6	2.50	22	94	78	9	4.50	⋯	43	76	57	49	73.50
-393	21	30	45	6	2.50	22	9	80	9	4.50	⋯	50	44	50	49	73.50
-392	20	85	50	6	2.50	23	27	77	9	4.50	⋯	45	35	55	49	73.50
-391	18	29	56	6	2.50	21	47	80	9	4.50	⋯	49	23	51	49	73.50
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮
-1	22	72	44	6	2.50	22	87	76	9	4.50	⋯	48	97	52	49	73.50
0	18	25	54	6	2.50	22	43	79	9	4.50	⋯	44	31	56	49	73.50

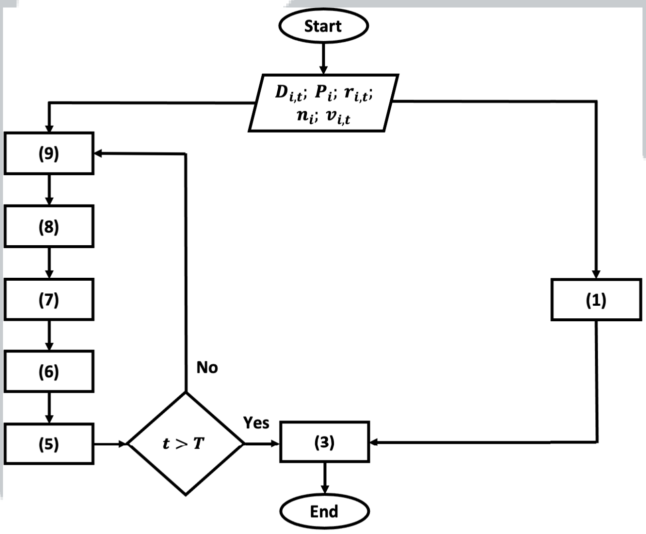

Fig. 2. Flowchart of reinforcement learning on the credit risk-based pricing.

III. SIMULATION

When the business begins, it is assumed that the credit company do not know the acceptance rate function $\bar{F}_i(r_i)$. The determined rates probably do not give maximum profitability for the company. Suppose after 400 batch of applicants, the company decides to apply reinforcement learning. Here we give a partial illustrative on implementation for this case. We use randomly generated dataset of size 400 as shown partially in Table I based on the cost of capital $r_c = 0.08$. The generated dataset follows the distribution of sample data collected from PT Amartha Mikro Fintek as one of financial technology peer-to-peer lending company in Indonesia during the period 2014 – 2020. For simplicity, we assume $v_{i,t} = 0$. The dataset consists of 27 customer segments based on their credit scores (grade A, B, and C), plafond (50, 100, and 150 million rupiahs), and tenor (1 year, 2 years, and 3 years).

We apply regression $y = a_i + b_i r + \epsilon$ on those 400 data to have initial parameters, $\hat{\theta}_{i,400} = \left(\begin{smallmatrix} \hat{\alpha}_{i,400} \\ \hat{\beta}_{i,400} \end{smallmatrix}\right)$ and covariance matrix $\sigma_i^2 R_{i,400}$. For instance, for the first segment we have $\hat{\theta}_{1,400} = \left(\begin{smallmatrix} 1.9999 \\ -9.9956 \end{smallmatrix}\right)$, and $R_{1,400} = \left(\begin{smallmatrix} 0.49368 & -2.45498 \\ -2.45498 & 12.27028 \end{smallmatrix}\right)$. After determining all initial parameter, the next step is to find the optimal profitability of each segment at $t = 401$ by:

$$\max_{r_{i,401} \geq 0} \hat{\pi}_{i,401}(r_{i,401}) = D_{i,401} \hat{\bar{F}}_{i,401}(r_{i,401}) \, \mathrm{PVNII}(P_i, r_{i,401}, n_i),$$

where $\hat{\bar{F}}_{i,401}(r_{i,401}) = \exp\left[-\left(\hat{\alpha}_{i,400} + \hat{\beta}_{i,400} r^*_{i,401}\right)\right]$.

After finding the optimal rates vector $r^*_{i,401} = (r^*_{1,401}, r^*_{2,401}, \cdots, r^*_{N,401})$, the Kalman Filter is then applied to update all parameters $\hat{\theta}_{i,400}$ to $\hat{\theta}_{i,401}$ as well as $R_{i,400}$ to $R_{i,401}$. For instance, the optimal rate for the first segment is $r^*_{1,401} = 0.24616$, giving predicted optimal profitability 145.18 million rupiahs. Based on Equation (7), (8), and (9), we obtain updated parameters $\hat{\theta}_{1,401} = \left(\begin{smallmatrix} 1.9939 \\ -9.9652 \end{smallmatrix}\right)$ and $R_{1,401} = \left(\begin{smallmatrix} 0.48178 & -2.39415 \\ -2.39415 & 11.95940 \end{smallmatrix}\right)$. This illustration shows the reinforcement learning incrementally updates the parameters of acceptance rate function for each segment (that are, $\hat{\alpha}_i$ and $\hat{\beta}_i$) so that the optimal loan rate can be determined. The computations for the rest of segments are done similarly using GNU Octave programming language. Figure 3 shows the distribution of optimal rate per segment.

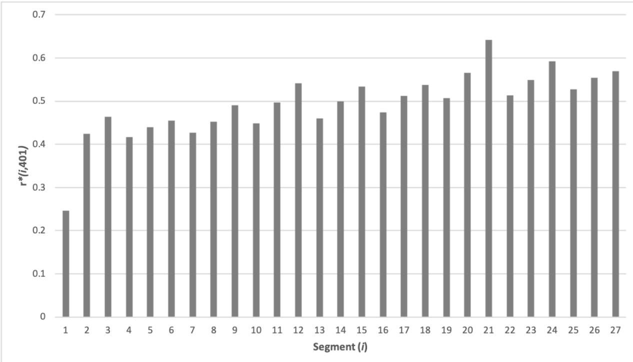

Fig. 3. Optimal rate per segment.

IV. Conclusion and Future Works

Reinforcement learning approach on the credit risk-based pricing models considered in this paper adjusts the parameters of acceptance rate function according to the current inputted data. This makes the model more flexible to the current situation compared to the regression method. The drawback of the model is that it is less realistic, since all segments are independent to each other. Nevertheless, this implies an easy calculation to find the optimal aggregate profitability. We suggest improvement to the model by considering additional constraints that represent relationship among segments. In addition, it is worth to consider converting discrete categorical segments into continuous numerical values that included in the model as an additional parameter.

Acknowledgment

This document is the results of the research project funded by Direktorat Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi / Badan Riset dan Inovasi Nasional (Grant No. 234/SP2H/LT/DRPM/2021).

References

[1] Oesterreichische Nationalbank and Financial Market Authority, “Credit Approval Process and Credit Risk Management: Guidelines on Credit Risk Management,” Vienna, Austria: Oesterreichische Nationalbank, 2004.

[2] Z. Ma, W. Hou and D. Zhang, “A credit risk assessment model of borrowers in P2P lending based on BP neural network,” PLoS ONE, vol. 16(8): e0255216, pp. 1–21, August 2021.

[3] C. Liu, H. Shi, Y. Cai, S. Shen and D. Lin, “A New Pricing Approach for SME Loans Issued by Commercial Banks on Credit Score Mapping and Archimedean Copula Simulation,” Journal of Business Economics and Management, vol. 20(4), pp. 618–632, 2019.

[4] S. Tran and P. Verhoeven, “Kelly Criterion for Optimal Credit Allocation,” Journal of Risk and Financial Management, vol. 14(434), pp. 1–15, September 2021.

[5] A. G. Walke, T. M. Fullerton Jr. and R. J. Tokle, “Risk-based loan pricing consequences for credit unions,” Journal of Empirical Finance, vol. 47, pp. 105–119, February 2018.

[6] W. Edelberg, “Risk-based pricing of interest rates for consumer loans,” Journal of Monetary Economics, vol. 53(8), pp. 2283–2298, November 2006.

[7] Y. Deng and S. Gabriel, “Risk-based pricing and the enhancement of mortgage credit availability among underserved and higher credit-risk populations,” Journal of Money, Credit and Banking, vol. 38(6), pp. 1431–1460, September 2006.

[8] F. Wang, “Risk-based pricing in high loan-to-value mortgage,” Munich Personal RePec Archive, pp. 1–27, September 2007.

[9] J. S. Spader, “Beyond disparate impact: Risk-based pricing and disparity in consumer credit history scores,” Review of Black Political Economy, vol. 37(2), pp. 61–78, January 2010.

[10] R. Phillips, “Optimizing prices for customer credit,” Columbia University Center for Pricing and Revenue Management Working Paper, vol. 2013–1, March 2013.

[11] Experian Decision Analytics, “Optimization in retail consumer lending,” An Experian Decision Analytics White Paper, 2013.

[12] D. Boyd, M. Gordon, J. Andersson, C. H. C. Tai, F. Yang, A. Kolamala, G. Cook, T. Guardino, M. Purang, P. Krishnamurty, M. Cooke, R. Nandiwada, B. Monteiro and S. Haas, “Target pricing system,” US Patent, No. 6963854 B1, 2005.

[13] R. Phillips, “Customized pricing,” The Oxford Handbook of Pricing Management, November 2012.

[14] H. Wang, “Credit Risk Management of Consumer Finance Based on Big Data,” Mobile Information Systems, vol. 2021: 8189255, pp. 1–10, 2021.

[15] N. Bussmann, P. Giudici, D. Marinelli and J. Papenbrock, “Explainable Machine Learning in Credit Risk Management,” Computational Economics, vol. 57, pp. 203–216, 2021.

[16] A. X. Carvalho and M. L. Puterman, “Dynamic pricing and reinforcement learning,” In Proceedings of the International Joint Conference on Neural Networks IEEE, pp. 2916–2921, July 2003.