Asian Journal of Economics and Banking

ISSN 2615-9821

http://ajeb.buh.edu.vn/Home

Posterior Summary of Bayes Error Using Monte-Carlo Sampling and Its Application in Credit Scoring

Ha Che-Ngoc*

Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Article Info

Received: 02/01/2020
Accepted: 16/3/2020
Available online: In Press

Keywords
Beta distribution, Posterior distribution, Bayes error, Monte-Carlo sampling

JEL classification

C15

MSC2020 classification

62H30

Abstract

Bayesian classifier is one of the data classification methods that are of interest. In the Bayesian classifier, Bayes error, Pe is an important measure because it can estimate the error of the model built through the calculation of the posterior probability function's overlapping area. The exact calculation of Pe depends on the exact calculation of likelihood functions and the prior probability of each type. In previous studies, the prior probability has been considered as a fixed value only, hence, the Bayes error is usually a fixed value. This sometimes leads to unreasonable results. To fill the mentioned research gap, this paper considers the prior probability q in Bayesian classifier as a distribution, and looks insight the posterior distribution of Bayes error, using Monte-Carlo simulation. Finally, the proposed method is applied to credit scoring data of a bank in Vietnam. Based on the results, we can determine whether the Bayesian classifier is suitable for data or not. In addition, the prior parameter setting can be tested through sensitivity analysis.

*Corresponding author: Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam. Email: chengocha@tdtu.edu.vn

1 INTRODUCTION

Classification or supervised learning is an important problem that is applied in almost all fields. In recent years, thanks to the rapid development of computers, supervised learning problems have been researched and developed extremely diverse. In comparison to other classification methods, the Bayesian classifier has some advantages, such as simplicity, explainability, etc. Therefore, along with the development of deep learning methods, Bayesian classifier has been widely used [1]. Another advantage of the Bayesian classifier is it's ability to assess the model's risk, by calculating Bayes error. Bayes error is a theoretical measure that can estimate the error of the model built through the posterior probability density function's overlapping area. The larger this overlap is, the higher the error of the Bayesian model is. Thus, we can estimate the error and evaluate the suitability of the Bayesian model without testing on cross-validation sets and test sets as other methods. This feature is very useful when we apply to problems with little data. On these problems, splitting the data into training sets, cross-validation and test sets will make the amount of data on the sets, especially the training set, which is not generalized enough to build a model. Therefore, Bayesian classifier and Bayes error have been topics of great interest in recent years. According to [5, 7], Bayes error depends on the correct determination of two factors: the probability density function (pdf) and the prior probability q. This implies that we can only accurately estimate the errors and the suitability of the Bayesian classifier if and only if the above two factors are reasonably determined, otherwise, the Bayes error will not reflect reality. So far, the problem of determining the pdf or the likelihood function has been considered by many researchers in both theory and practice [5]; therefore, in this study, we only apply the results of the existing methods. When the pdfs have been well defined, determining the prior probability q is an important factor to improve the performance of Bayesian classifier. However, this issue still depends on the experience and opinion of researchers too much.

There are many methods for setting prior probability. For example, we can choose the prior probability based on equal priors ($q_1 = q_2 = 1/2$), based on the training set $q_i = N_i/N$, or based on the Laplace formula $q_i = (N_i + 1)/(n + N)$, where N_i is the number of elements in w_i, n is the number of dimensions and N is the number of observations in training data. When we choose different prior probability values, the Bayes error received will be different. This shows the irrationality of the above methods when using fixed prior probability value and output a specific value of Bayes error. The irrationality will be higher when the data set becomes larger and has higher fluctuation. Some references using fixed values of prior probabilities for calculating Bayes error can be listed as [2, 4, 7, 9, 10].

Besides, in some recent studies, the prior probability q has been considered as a random variable with Beta distribution [6, 11, 14]. The prior and posterior distributions of $|1 - 2q|$ and $q/(1 - q)$ are also clarified in theory and practical application in these studies. It can be said that "liberating" the view, studying q as a distribution instead of a specific value is a significant step forward. However, when classifying and calculating Bayes errors in actual data, the studies selected a fixed value derived from the posterior distribution of q, such as mean or mode. Although these values have been chosen so that they can well represent the posterior distribution, choosing a single value of q may result in the same limitations as the previous methods.

In order to fill the above-researched gap, in this paper, we examines the prior probability q as a distribution; however, in the Bayes error calculation, we do not choose a specific value of q to represent the posterior distribution $\{q|Y = y\}$, instead we try to keep all the information from this posterior distribution (where y is the number of observations labeled as w_1 in training data). Particularly, we simulate N_1 samples q having the distribution of $\{q|Y = y\}$ by Monte-Carlo method, thereby computing N_1 corresponding Bayes error Pe values. Based on this simulation, for the first time, we can look insight the Bayes error distribution rather than a single value or bounds as before. Some other posterior inferences such as mean, variance, credible interval, also can be made. Compared to previous studies that only provided a point estimation, the simulation results are expected to provide more information about Bayes error. This can lead to a more comprehensive model evaluating.

In recent years, Vietnam's financial market has grown strongly and banks have had more opportunities and challenges from their credit activities. In the banking industry, credit scoring is an important tool that can determine the client's ability to repay the debt. If the lending is too easy, the bank may have bad debt problems. In contrast, the bank will miss a good business. To classify the ability to repay bank debt, various kinds of Bayesian classifier have recently been studied in Vietnam [8,12,13]. However, credit data are diverse, volatile and uncertain, hence, it would be better if we can pre-evaluate the model's suitability through Bayes error before applying Bayesian classifier. Therefore, the proposed Bayes error calculation approach will be used as a criterion to show the degree of suitability of Bayesian classifier. We compare the proposed method with the Bayes error calculated with a fixed value of q, and evaluate the effectiveness of the methods based on the empirical errors from the test set. Finally, the sensitivity analysis is performed to examine how the advanced distribution choice for q would affect the posterior information of Bayes error.

The remainder of this paper is organized as follows: Section 2 and Section 3 present relevant knowledge and the proposed framework. In Section 4, the proposed frame work is applied to credit scoring data in Vietnam. Finally, Section 5 is the conclusion.

2 BAYESIAN CLASSIFIER AND BAYES ERROR

2.1 Bayesian Classifier

We consider k classes, $w_1, w_2, \ldots, w_k$, with the prior probability q_i, $i = \overline{1,k}$. $\mathbf{X} = \{X_1, \ldots, X_n\}$ is the n-dimensional continuous data with $\mathbf{x} = \{x_1, \ldots, x_n\}$ is a specific sample. According to [5,7], a new observation $\mathbf{x}_0$ belongs to the class w_i if

$$P(w_i|\mathbf{x}) > P(w_j|\mathbf{x}) \quad \text{for} \quad 1 \le j \le k, j \ne i. \tag{1}$$

In continuous case, $P(w_i|\mathbf{x})$ is calculated by:

$$P(w_i|\mathbf{x}) = \frac{P(w_i) f(\mathbf{x}|w_i)}{\sum_{i=1}^{k} P(w_i) f(\mathbf{x}|w_i)} = \frac{q_i f_i(\mathbf{x})}{f(\mathbf{x})}. \tag{2}$$

Because $f(\mathbf{x})$ is the same for all classes, the classification's rule is:

$$q_i f_i(\mathbf{x}) > q_j f_j(\mathbf{x}) \Leftrightarrow \frac{f_i(\mathbf{x})}{f_j(\mathbf{x})} > \frac{q_j}{q_i}, \forall i \ne j, \tag{3}$$

where q_i and $f_i(\mathbf{x})$ are the prior probability and the probability density function of class i, respectively.

For the binary classification investigated in this paper, the new observation $\mathbf{x}_0$ belongs to the class w_1 if $q_1 f_1(\mathbf{x}_0) > q_2 f_2(\mathbf{x}_0)$ or $P(w_1|\mathbf{x}) > 0.5$ and vice versa.

2.2 Bayes Error

Let $g_i(\mathbf{x}) = q_i f_i(\mathbf{x})$, $g_{\max}(\mathbf{x}) = \max_i \{g_i(\mathbf{x})\}$ and $g_{\min}(\mathbf{x}) = \min_i \{g_i(\mathbf{x})\}$. According to [5,7], Bayes error is calculated by Formula 4,

$$Pe_{1,2,\ldots,k}^{(q)} = 1 - \int_{R^n} g_{\max}(\mathbf{x})\, \mathrm{d}\mathbf{x} \tag{4}$$

where Pe is Bayes error, n is the dimension of the data.

2.3 Bayes Error in Binary Classification

In case of binary classification,

$$Pe = 1 - \int_{R^n} g_{\max}(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{R^n} g_{\min}(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{5}$$

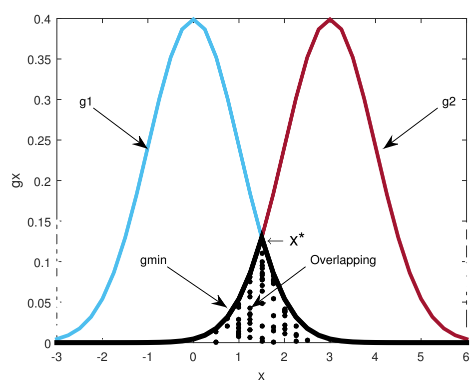


Fig. 1: Bayes error in case of binary classification

To understand Bayes error in the case of binary classification, Figure 2.3 illustrates a commonly used Bayesian classifier, which has one dimension. Assuming that we have two classes consisting of w_1 and w_2, and $\mathbf{x}$ is the vector of independent variables (in fact, $\mathbf{x}$ is a scalar in univariate model). The probability of a false prediction is calculated by:

$$Pe = Pe_1 + Pe_2 \tag{6}$$

where Pe_1 is the probability of that the predicted class is 1 but the actual class is 2, and Pe_2 is the probability of that the predicted class is 2 but the actual class is 1. Figure 2.3 shows that

$$Pe_1 = \int_{-\infty}^{\mathbf{x}*} g_2(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{-\infty}^{\mathbf{x}*} g_{\min}(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{7}$$

and

$$Pe_2 = \int_{\mathbf{x}*}^{+\infty} g_1(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{x}*}^{+\infty} g_{\min}(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{8}$$

with $\mathbf{x}*$ is root of equation $g_1(\mathbf{x}) = g_2(\mathbf{x})$.
Combining (6), (7) and (8), we obtain:

$$Pe = \int_{-\infty}^{+\infty} g_{\min}(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{9}$$

The Bayes error calculated by Formula (9) can be visually interpreted as the area of the overlapping region between g_1 and g_2. Using the Bayes error, we can estimate the probability of an incorrect prediction without performing on a cross-validation set. This property is an advantage of Bayesian classifier in comparison to other methods, and can be fully applicable to any classification problem.

In the case of univariate normal distribution, we can find out the specific expression for $g_{\min}(\mathbf{x})$ [5, 12]; hence, we can also identify Pe. In the case of arbitrary multivariate distributions, the specific expression of $g_{\min}$ are difficult to be identified; hence, Quasi Monte-Carlo method is applied to approximate the value of integrals.

Quasi Monte-Carlo approximation
Let $Pe = \int_{-\infty}^{+\infty} g_{\min}(\mathbf{x})\mathrm{d}\mathbf{x}$ is the Bayes error that needs to be computed. The Quasi Monte-Carlo approximate Pe by:

$$\hat{P}e = \frac{\sum_{i=1}^{N_2} g_{\min}(\mathbf{x}_i)}{N_2}\mathrm{Mes}(A) \tag{10}$$

where $\mathbf{x}_i$ random points sampled in space A, N_2 is the sample size, $\mathrm{Mes}(A)$ is the measure of A which is often equal to 1 when data are standardized.

2.4 Posterior Distributions under Beta Prior Distributions

A random variable q, which is between 0 and 1, has a beta(a, b) distribution if

$$f(q) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} q^{a-1}(1-q)^{b-1} \text{ for } 0 \le q \le 1, \tag{11}$$

where $f(q)$ is the density function of q and $\Gamma(a) = \int_0^{\infty} x^{a-1}e^{-x}\mathrm{d}x$. Now let $Y \sim \text{binomial}(N, q)$ with $q \sim \text{beta}(a, b)$, according to [3], we have

$$\{q|Y = y\} \sim \text{beta}(a + y, b + N - y), \tag{12}$$

where y is the number of interesting events occurs after N trials.

3 PROPOSED FRAMEWORK

For filling the researched gap in literature which utilized a fixed prior probability when computing the Bayes error, this section proposed a new framework to approximate the Bayes error distribution using Monte-Carlo simulation. In particular, the prior probability q is investigated under Beta prior and updated via training data set to receive its posterior distribution $\{q|Y = y\}$.

We then simulate N_1 values of q using its posterior distribution. For each value q_i, $i = 1, \ldots, N_1$, we simulate N_2 n-dimensional points $\mathbf{z}$ with $Z_{ijk} \sim$ U$(0,1), j = 1, \ldots, N_2$, $k = 1, \ldots, n$ (the data have been standardized); compute $\min\{(g_1|q, \mathbf{z}_{ij}), (g_2|q, \mathbf{z}_{ij})\}$ that is the value of $g_{\min}$ at $\mathbf{z}_{ij}$, $j = 1, \ldots, N_2$; compute the $\hat{Pe}_i$ using Quasi Monte-Carlo method. Finally, we obtain N_1 values of $\hat{Pe}_i$ that can be used for the posterior inferences, such as computing the mean, estimating the credible interval, approximating the distribution, etc. In short, let beta(a, b) is the prior distribution of q and y is the number of observations belonging to w_1 in N observations in training set, the proposed approach is summarized as the table follows.

Using the obtained $\{Pe_1, Pe_2, \ldots, Pe_{N1}\}$, we can approximate:

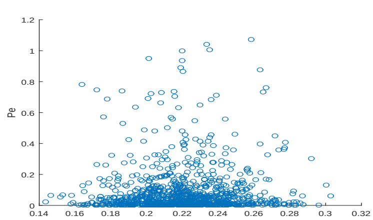
- $\mathrm{E}\left[Pe|y\right] \approx \overline{Pe} = \sum_{i=1}^{N_1} \frac{Pe_i}{N_1}$.
- $\mathrm{Var}\left[Pe|y\right] \approx \sum_{i=1}^{N_1} \frac{(Pe_i - \overline{Pe})^2}{N_1 - 1}$.
- $f(Pe|y) \cong$ the empirical distribution of $\{Pe_1, Pe_2, \ldots, Pe_{N_1}\}$.
- Etc.

4 APPLICATION TO CREDIT SCORING

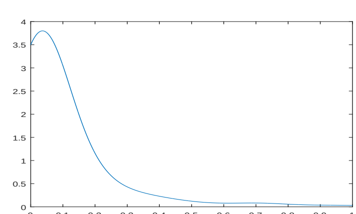
In this section, we apply the proposed framework to evaluate the suitability of the Bayesian classifier in Vietcombank's customer in Vietnam. The customers are companies that operate in important fields such as agriculture, industry, and commerce in Can Tho city. This data is provided by responsible organizations and has been studied by [13]. In the original dataset, there are 13 independent variables and one dependent variable consisting of two

Sample $q_1 \sim \mathrm{beta}(a+y, b+N-y)$	sample $\mathbf{z}_{11} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{11}\}$ sample $\mathbf{z}_{12} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{12}\}$ ... sample $\mathbf{z}_{1N_2} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{1N_2}\}$ compute Pe_1
Sample $q_2 \sim \mathrm{beta}(a+y, b+N-y)$	sample $\mathbf{z}_{21} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{21}\}$ sample $\mathbf{z}_{22} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{22}\}$ ... sample $\mathbf{z}_{2N_2} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{2N_2}\}$ compute Pe_2
...	
Sample $q_{N_1} \sim \mathrm{beta}(a+y, b+N-y)$	sample $\mathbf{z}_{N_11} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{N_11}\}$ sample $\mathbf{z}_{N_12} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{N_12}\}$ ... sample $\mathbf{z}_{N_1N_2} \sim \mathrm{U}(0,1)$, compute $\{g_{\min}\vert q_1, \mathbf{z}_{N_1N_2}\}$ compute Pe_{N_1}

classes: w_1-bad debt and w_2-good debt. However, according to [13], the three variables including Financial, Interest, and Profits result in better performance than other variables. Therefore, the reduced data set consisting of 214 companies, three independent variables, and one dependent variable will be used in this paper. This data set is divided into two parts: training and testing with a ratio of 7:3 to evaluate the effectiveness of the methods. For the proposed method, according to sources in the media in Vietnam, the bad debt ratio of banks is no more than 2%, so we choose distribution of prior probability q is beta$(a = 2, b = 98)$ so that $\mathrm{E}[q] = 0.02$ and $a + b$ are not too big because of belief in type reports is not high. We continue to use the suggested process in Section 3, with $N_1 = N_2 = 1000$, simulation results of Pe and q, experimental distribution of Pe received after the simulation is shown in Figure 4



(a) Scatter plot of simulated q and Pe



(b) Empirical distribution of Pe

Fig. 2: Some results of q and Pe.

Figure 4a shows the scatter plot of simulated q and Pe. Recall we assumed that $q \sim$ beta$(a = 2, b = 98)$. Since training data set (70% of data) contains 53 cases of bad debt and 150 cases of good debt, the corresponding sample information, posterior distribution, and posterior mean of q are $\{y = 53, N = 150\}$ and $q \sim$ beta$(a+y = 55, b + N - y = 195)$, $\mathrm{E}[q|y = 55] = 55/(55 + 195) = 0.22$. It can be seen from Figure 4a that the simulated values of q are concentrated around this area. For Pe, most simulation values are smaller than 0.2, however, there are also many cases where Pe receives large values, even greater than 0.6 or close to 1. The experimental distribution of Pe is also shown in Figure 4b. We see that this distribution seems to be similar to Beta distribution, which needs to be clarified in further studies. From the values of Pe, we can easily approximate characteristic parameters such as mean, posterior credible interval, HDD. The results of the proposed method and the other ones are also shown in Table 1.

In Table 1., Bayesian method with prior probability calculated by equal priors, training data set, Laplace method are respectively called BayesU, BayesT, BayesL; BayesR and BayesD are proposed by [14], mean for posterior distribution of $q/(1-q)$ and $|1-2q|$ are representative values, BayesM is the proposed method.The results from Table 1. show that most of the methods produce fairly small Bayes error values (most are less than 16 %), this result initially shows that the Bayesian classifier seems appropriate for the credit

Table 1. Bayes error of different methods

	Bayes error	Variance	CI	HDD
BayesT	0.1547	-	-	-
BayesU	0.1340	-	-	-
BayesD	0.0123	-	-	-
BayesR	0.1236	-	-	-
BayesM	0.0921	0.0236	$[0.0003, 0.3951]$	$[0.0000, 0.2492]$

scoring application. However, the simulation results from the BayesM method presented above, as well as the results at the last line in Table 1. show the great variability of Bayes error Pe. In Figure 4, some simulated Pe values are very large and close to 1. The results in Table 1. also show that the upper bound of the 90% credible interval, CI, of Pe is up to nearly 0.4; this result will be even greater if we calculate the 95% and 99% CI of Pe. It can be implied that although the point estimates are quite small, there is still a high possibility of a miss-classification. Therefore, researchers have to examine the model more carefully or conduct further experiments before applying Bayesian classifier to the credit scoring problem. This means we will be able to experiment carefully with the proposed Bayes error calculation. It also an advantage of the proposed method compared to previous studies.

Finally, we conducted a sensitivity analysis to test how the choice of a prior probability would affect Bayes errors and accuracy. As mentioned above, according to the sources in the media in Vietnam, the bad debt ratio of banks is not more than 2 %, so initially we chose $q \sim \text{beta}(a = 2, b = 98)$ so that $\mathrm{E}[q] = 0.02$ and $a+b$ are not too big because the confidence in reports of this type is not high. In this section, we will examine q of the form $\text{beta}(a = 2 * k, b = 98 * k)$ with $k = 1, \ldots, 50$. The method set above still ensures $\mathrm{E}[q] = 0.02$ however the magnitude of $a + b$ will increase as k increases, showing stronger belief in the prior information. We also calculated the actual error received on 30% test data. The results obtained are shown in Figure 4

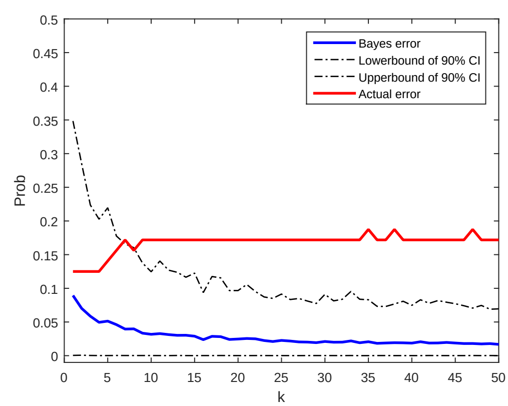


Fig.3: Result of sensitivity analysis

From Figure 4, we find that the Bayes error and the estimated credible interval width becomes smaller when k increases. This is natural because a greater value of k indicates a stronger belief in the prior information, which leads to a decrease in volatility and uncertainty (decrease in the width of the credible interval). However, Figure 4 also shows that the prior information we use is not accurate. Obviously, at low values of k, although the prior information is not accurate, the posterior information of Pe has been better adjusted through the sample collection process on the training set. As k increases, the effect of the prior information is larger and the effect of the sampling process is smaller. As a result, the estimated value of the Bayes error and the credible interval are not consistent with the actual error. Therefore, it would be better if we set prior information sources with a small confidence scale ($a + b$ small) as the original setting of the article.

5 CONCLUSION

In Bayesian classifier, the exact calculation of Bayes error, Pe, depends on the exact calculation of prior probability q. In previous studies, the prior probability has been considered as a fixed value only, hence, the Bayes error is usually a fixed value. This sometimes leads to unreasonable results. In this paper, the prior probability has been chosen based on Beta distribution, from which, the empirical distribution of Pe has been introduced through the Monte-Carlo simulation process, for the first time. The application in the credit scoring problem has shown that the proposed Bayes error calculation can achieve more information than previous studies. Specifically, we can estimate the distribution of Bayes error, $\{Pe|Y = y\}$, and its posterior summary, such as average, variance, CI, HDD, etc., and indicates a high possibility of a miss-classification when applying Bayesian classifier to the credit scoring problem. Finally, sensitivity analysis is applied to show the Bayes error sensitiveness to the prior information used. In the future, we will consider the distribution of q in the multi-classification problem and provide more comprehensive assessments in Bayesian classifier with Bayes error.

References

[1] Chollet, F., & Allaire, J. (2017). *Deep Learning with R.* Manning Publications.

[2] Ciabattoni, L., Cimini, G., Ferracuti, F., Grisostomi, M., Ippoliti, G., & Pirro, M. (2015). Bayes error based feature selection: An electric motors fault detection case study. In *Iecon 2015-41st annual conference of the ieee. industrial electronics society* 003893-003898.

[3] Hoff, P. D. (2009). *A first course in bayesian statistical methods* (Vol. 580). Springer.

[4] Moon, K. R., Hero, A. O., & Delouille, B. V. (2015). Meta learning of bounds on the bayes classifier error. In *2015 ieee signal processing and signal processing education workshop (sp/spe)* 13-18.

[5] Nguyen-Trang, T., & Vo-Van, T. (2017). A new approach for determining the

prior probabilities in the classification problem by Bayesian method. *Advances in Data Analysis and Classification*, 11 (3), 629-643.

[6] Pham-Gia, T., Turkkan, N., & Bekker, A. (2007). Bounds for the Bayes error in classification: a Bayesian approach using discriminant analysis. *Statistical Methods and Applications, 16*(1), 7-26.

[7] Pham-Gia, T., Turkkan, N., & Vo-Van, T. (2008). Statistical discrimination analysis using the maximum function. *Communications in Statistics Simulation and Computation, 37*(2), 320-336.

[8] Pizzo, A., Teyssere, P., & Vu-Hoang, L. (2018). Boosted Gaussian Bayes Classifier and its application in bank credit scoring. *Journal of Advanced Engineering and Computation, 2*(2), 131-138.

[9] Rajanna, R., Cao, L., & Viswanathan, R. (2016). Divergence and bayes error based soft decision for decentralized signal detection of correlated sensor data. In *2016 annual conference on information science and systems (ciss)* 490-495.

[10] Sekeh, S. Y., Oselio, B., & Hero, A. O. (2018). Multi-class bayes error estimation with a global minimal spanning tree. In *2018 56th annual allerton conference on communication, control, and computing (allerton)*, 676-681.

[11] Vo-Van, T. (2017a). Classifying by Bayesian Method and Some Applications. In *Bayesian inference* 39-61). InTech.

[12] Vo-Van, T. (2017b, feb). L1-distance and classification problem by Bayesian method. *Journal of Applied Statistics, 44*(3), 385-401. Retrieved from doi: 10.1080/02664763.2016.1174194

[13] Vo-Van, T., Che-Ngoc, H., & Nguyen-Trang, T. (2016). The prior probability in classifying two populations by Bayesian method. In *1st international conference on applied mathematics in engineering and reliability*, 35–40.

[14] Vo-Van, T., Tran-Phuoc, L., & Che-Ngoc, H. (2019). Classifying two populations by bayesian method and applications. *Communications in Mathematics and Statistics, 7*(2), 141-161.