Research on Personal Credit Scoring Based on Histogram of Oriented Gradient

Fan Liao*, Shu Huang, and Wu Jian

Abstract—**The topic of personal credit scoring is a prerequisite for the development of personal credit business and has become an important research area in the field of financial risk management. However, little attention has been paid to the field of personal credit scoring research on how to introduce additional features in improving personal credit scoring services. To solve this problem, a new method based on Histogram of Gradients Oriented (HOG) for personal credit scoring is proposed for enhancing the scoring capability. The proposed method utilizes the gradient relationship between different features of personal credit data and introduces additional features on the basis of the original features of personal credit data to construct a set of personal credit datasets based on HOG. The experimental results show that the personal credit dataset based on HOG not only has higher values but also better stability in terms of the four metrics (accuracy, recall, precision, and F1-value) for personal credit scoring when compared to the original personal credit dataset. Therefore, the proposed personal credit scoring method has been demonstrated to be both reliable and feasible.**

Index Terms—**Histogram of Oriented Gradient, personal credit scoring, credit dataset, additional features.**

I. Introduction

THE lending business has always held significant importance in financial activities and serves as the primary source of income for financial institutions. However, every loan that a financial institution provides is exposed to varying degrees of risk, making the quality of the lending business a crucial aspect for these institutions. The financial industry has devoted considerable effort to studying methods for assessing loan risks for an extended period, aiming to decrease the potential for loan defaults. In 1975, the implementation of the Equal Credit Opportunity Act established a legal framework in the United States for granting personal loans based on individuals' credit scores.

The topic of personal credit scoring has gained increasing attention in recent years. In 2000, West [1] incorporated neural networks into credit scoring models and compared their effectiveness to conventional methods such as Linear Discriminant Analysis (LDA) algorithm. The credit scoring model established on neural networks yields more precise results. In 2002, Lee et al. [2] introduced a two-stage hybrid credit scoring model that combines discriminant analysis and neural networks methods. The neural networks construct a more robust statistical foundation with the aid of discriminant analysis, resulting in scoring outcomes that are more precise than those generated by a credit scoring model built solely on neural networks. In 2003, Baesens et al. [3] pointed out that research in the field of credit scoring has few analytical tools, mostly focusing on how to make the scoring results more accurate, proposing the use of neural networks rule extraction and decision tables for credit scoring of credit datasets, and experimentally proving the effectiveness of neural networks rule extraction applied to the field of credit scoring. In 2005, Ong et al. [4] utilized genetic programming in constructing credit scoring models, verifying the models through the German and Australian credit datasets. The experimental results demonstrate the superiority of the genetic programming based model's scoring outcomes in comparison to those from a neural network-based model. In 2007, Huang et al. [5] put forward a model for hybrid credit scoring that utilizes Support Vector Machine (SVM) and genetic algorithms to enhance the accuracy of feature selection in credit datasets. In 2008, Mavri et al. [6] introduced a two-stage dynamic credit scoring model for evaluating credit applicants. They confirmed the model's accuracy using data from a European bank, which can serve as a basis for bank's credit operations. In 2008, Yu et al. [7] proposed a credit risk assessment model based on the integration of a six-stage neural networks that performs credit scoring with a high accuracy. In 2009, Antonakis et al. [8] suggested that the Naive Bayes (NB) algorithm is frequently employed in credit scoring tasks because of its easy-to-understand nature; however its results are typically less granular when applied to large datasets. Thus, there are limited situations in which the NB algorithm is appropriate for credit scoring capacity. In 2009, Zhao et al. [9] presented a credit scoring model for assessing the creditworthiness of applicants, which was implemented by a Hong Kong-based Chinese bank to analyze credit applicant repayment risk. In 2009, Zhou et al. [10] proposed a credit scoring model that utilizes SVM and the Area Under Curve (AUC) metric. The accuracy of the model was confirmed by using data from a Hong Kong-based bank in China. This method can increase profits for the bank while ensuring organized credit control. Furthermore, Zhou et al. [10] applied their credit scoring model to German and Australian credit datasets, successfully verifying the model's scoring capability. The experimental results prove that the credit scoring model combining SVM and AUC algorithms scores better than traditional credit scoring models such as Linear Regression (LR) and Decision Trees (DT) algorithms.

Manuscript received December 30, 2024; revised April 14, 2025.

This work was supported in part by the Scientiffc Research Foundation for High-level Talents in Nanjing Vocational College of Information Technology (YB202415), and in part by the 2023 Jiangsu Province Service Customization Network Application Engineering Research Center (Innovation and High tech Development Department of Jiangsu Provincial Development and Reform Commission (2023) No. 520).

Fan Liao is a lecturer of School of Network and Communication, Nanjing Vocational College of Information Technology, Nanjing 210023, P.R. China (e-mail: liaofan111@qq.com).

Shu Huang is a professor of Laboratory of Computer Science and Technology, Southeast University, Nanjing 210096, P.R. China (e-mail: shuhuangseu@hotmail.com).

Wu Jian is an associate professor of Laboratory of Computer Science and Technology, Southeast University, Nanjing 210096, P.R. China (e-mail: wujianseu@hotmail.com).

In 2011, Chuang et al. [11] proposed a two-stage credit scoring model based on case-based reasoning and verified the model's accuracy using the German credit dataset, which searches for incorrectly categorized credit cases to improve the accuracy of credit scoring. In 2016, Wei et al. [12] conducted a study on the influence of social network data on credit scoring. In 2018, Kapoor et al. [13] suggested that Social Media provides a significant amount of data that can be utilized to assess an individual's credit risk; therefore, it is a vital source of information for credit risk assessment. In 2020, Fu et al. [14] proposed a two-stage credit scoring model based on neural networks and bidirectional long-term and short-term memories and verified the model accuracy using a real dataset of 1,000 online lending platforms, which improves the accuracy of credit scoring. In 2020, Zhang et al. [15] proposed a multi-stage hybrid credit scoring model based on stacked integrated learning, which improves the accuracy of credit scoring results. In 2023, Li et al. [16] proposed a federated learning credit scoring model based on utilizing data from multiple institutions, which can be advantageous for small and medium-sized banks and other organizations with limited data samples. In 2023, Zhou et al. [17] proposed a KCSMOTE credit scoring model based on utilizing a few classes of clustering centers as the base point for enhancing data balancing, which appears to be more effective in handling unbalanced fraud data. In 2024, Talaat et al. [18] proposed a new interpretable credit scoring model based on a combination of deep learning and interpretable AI techniques, which improves the accuracy of credit card default predictions. In 2024, Bennehalli et al. [19] proposed a robust credit scoring model based on a one-sided selection methodology in conjunction with information gain and Pearson correlation, which improves the accuracy of credit in both balanced and unbalanced datasets. In 2024, Chen et al. [20] proposed a fiscal risk controlling model based on a data structure and random subspace, which contributes to the management of risk in the context of Internet finance. In 2024, Idrees et al. [21] proposed a solution to the challenge posed by unbalanced credit card datasets, achieving a balanced dataset, thereby improving the efficiency of the system designed for the detection of fraudulent activity. In 2024, Chiang et al. [22] proposed an inventory system with present value and credit period, which serves to enhance the existing body of knowledge surrounding future cash outflows within the context of financial credit environment inventories and models. In 2025, Meng et al. [23] proposed a novel approach for short-term personal credit load forecasting, which improves personal credit load forecasting performance while maintaining personal credit data security.

In summary, research in the area of personal credit scoring has undoubtedly focused on how to build a better personal credit scoring model, and little attention has been paid to the introduction of additional features in the credit scoring process.

In this paper, we propose a new personal credit scoring method that utilizes the gradient relationship between different features of personal credit data and introduces additional features on the basis of the original features of personal credit data to construct a set of personal credit datasets based on Histogram of Oriented Gradient (HOG) [20].

II. Methods

Personal credit data x of size $m \times n$ is given as follows

$$x \in R^{m\times n} \tag{1}$$

Equation (1) defines x as a digital signal containing personal credit data, $R^{m\times n}$ as a $m \times n$ matrix $(1 \leq m, 2 \leq n)$, m as the number of items of personal credit data, and n as the number of variables (which can be either numeric or categorical) contained in each item of personal credit data.

The last variable of personal credit data x is given as follows

$$x_{i,n} = \begin{cases} 1, good\ credit \\ 2, bad\ credit \end{cases}, (1 \leq i \leq m) \tag{2}$$

where $x_{i,n} = 1$ represents good credit and $x_{i,n} = 2$ represents bad credit.

Here, we temporarily set aside the last variable of the personal credit data x, so the personal credit data y of size $m \times (n-1)$ is given below

$$y_{i,j} = x_{i,j}, (1 \leq i \leq m, 1 \leq j \leq n-1) \tag{3}$$

Equation (3) defines y as a $m \times (n-1)$ matrix, i as the rows where the personal credit data y and x are located, j as the columns where the personal credit data y and x are located, $y_{i,j}$ as the row i and column j of the personal credit data y, and $x_{i,j}$ as the row i and column j of the personal credit data x.

In order to greatly reduce the impact on the results caused by the large differences in the values of the variables in different columns of the personal credit data y, the personal credit data y needs to be normalized by columns.

We now introduce the row vector $y^{max} \in R^{1\times(n-1)}$, which is defined as

$$y_j^{max} = max(y_{i,j}), (1 \leq i \leq m, 1 \leq j \leq n-1) \tag{4}$$

where y_j^{max} represents the value located in row 1 and column j of matrix y_j^{max}, and it denotes the maximum value in the variable in column j of the personal credit data y.

So, we can divide each variable in Equation (4) by the maximum value of the variable data in that column.

The mathematical definition of $y_{i,j}^{normalization}$ of size $m \times (n-1)$ is as follows

$$y_{i,j}^{normalization} = \frac{y_{i,j}}{y_j^{max} + \varepsilon} \tag{5}$$

where the positive parameter ε is a constant that is very close to zero in order to avoid having a zero denominator.

The value of Equation (5) has been obtained. In order to incorporate additional features on the basis of the original features of the personal credit data , it is also necessary to construct a three-dimensional matrix $CELL$ of size $e \times f \times m$, which is given as follows

$$CELL_{k,h,g} = y_{i,j}^{normalization} \tag{6}$$

Equation (6) defines $CELL_{k,h,g}$ as a three-dimensional matrix with e rows, f columns, and g pages $(1 \leq e \leq n-1$, $1 \leq f \leq n-1$ and $1 \leq g \leq m)$.

Here, parameters e, f, n, k, h, j, g, and i need to share the following relations

$$e \times f = n - 1$$

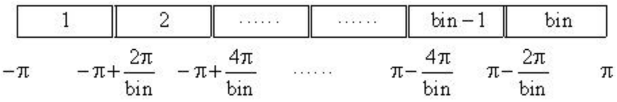


Fig. 1. The HOG of bin equal parts

$$(k-1)\times f+h=j$$

$$g=i$$

Therefore, the gradient of the $CELL_{k,h,g}$ matrix for personal credit data in the row direction can be defined as

$$CELL^{line}_{k,h,g}=\begin{cases} CELL_{k+1,h,g}-CELL_{k,h,g}, \\ (1\le k\le e-1, 1\le h\le f, 1\le g\le m) \\ 0-CELL_{k,h,g}, \\ (k=e, 1\le h\le f, 1\le g\le m) \end{cases} \tag{7}$$

And the gradient of the $CELL_{k,h,g}$ matrix for personal credit data in the column direction can be defined as

$$CELL^{column}_{k,h,g}=\begin{cases} CELL_{k,h,g}-CELL_{k,h+1,g}, \\ (1\le k\le e, 1\le h\le f-1, 1\le g\le m) \\ 0-CELL_{k,h,g}, \\ (1\le k\le e, h=f, 1\le g\le m) \end{cases} \tag{8}$$

We now calculate the gradient magnitude of the $CELL_{k,h,g}$ matrix for personal credit data, which is defined as

$$CELL^{magnitude}_{k,h,g}=\sqrt{(CELL^{line}_{k,h,g})^2+(CELL^{column}_{k,h,g})^2} \tag{9}$$

where $1\le k\le e$, $1\le h\le f$ and $1\le g\le m$.

Let $\theta_{k,h,g}$ be the gradient direction of the $CELL_{k,h,g}$ matrix for personal credit data. Therefore, $\theta_{k,h,g}$ is given as follows

$$\theta_{k,h,g}=\begin{cases} arctan(\frac{CELL^{column}_{k,h,g}}{CELL^{\text{line}}_{k,h,g}}), \\ (CELL^{line}_{k,h,g}>0) \\ arctan(\frac{CELL^{column}_{k,h,g}}{CELL^{\text{line}}_{k,h,g}})+\pi, \\ (CELL^{column}_{k,h,g}\ge 0, CELL^{\text{line}}_{k,h,g}<0) \\ arctan(\frac{CELL^{column}_{k,h,g}}{CELL^{\text{line}}_{k,h,g}})-\pi, \\ (CELL^{column}_{k,h,g}<0, CELL^{\text{line}}_{k,h,g}<0) \\ \frac{\pi}{2}, \\ (CELL^{col}_{k,h,g}\ge 0, CELL^{line}_{k,h,g}=0) \\ -\frac{\pi}{2}, \\ (CELL^{column}_{k,h,g}<0, CELL^{line}_{k,h,g}=0) \end{cases} \tag{10}$$

The subsequent step is to generate a HOG, as illustrated in Figure1.

Figure 1 shows that the range from $-\pi$ to $-\pi$ is divided into equal bin parts (bin $\ge$ 2). The part 1 covers the range $\left[-\pi,-\pi+\frac{2\pi}{\text{bin}}\right)$, the part 2 covers the range $\left[-\pi+\frac{2\pi}{\text{bin}},-\pi+\frac{4\pi}{\text{bin}}\right)$, ..., the part bin-1 covers the range $\left[\pi-\frac{4\pi}{\text{bin}},\pi-\frac{2\pi}{\text{bin}}\right)$, and the part bin covers the range $\left[\pi-\frac{2\pi}{\text{bin}},\pi\right)$.

The gradient magnitude and gradient direction of the $CELL_{k,h,g}$ matrix for personal credit data are obtained in Equations (9) and (10). In order to expand the additional features of personal credit data, we also need to accumulate the gradient magnitude into bin equal intervals according to the gradient direction. Thus, we create the matrix $y^{bin}\in R^{m\times bin}$ used to refer to the additional features of personal credit data, which is defined as

$$y^{bin}_{i,j}=\sum_{h=1}^{f}\sum_{k=1}^{e}CELL^{magnitude}_{k,h,g} \tag{11}$$

where $1\le i\le m$, $1\le j\le$ bin , $i=g$ and $\theta_{k,h,g}\in\left[-\pi+\frac{(j-1)\times 2\pi}{bin},-\pi+\frac{j\times 2\pi}{bin}\right)$.

Introducing the additional features y^{bin} into the personal credit data x, we obtain the data x^{bin} as follows

$$x^{bin}_{i,j}=\begin{cases} x_{i,j}, 1\le j\le n-1 \\ y^{bin}_{i,j-n+1}, n\le j\le n+bin-1 \\ x_{i,n}, j=n+bin \end{cases} \tag{12}$$

where $x\in R^{m\times(n+bin)}$.

III. Conclusions and Future Work

The data utilized in this study was retrieved from the German credit dataset (http://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)) offered by the University of California, Irvine, comprising of 1,000 credit data entries.

The mathematical definition of accuracy is as follows

$$Accuracy=\frac{TP+TN}{TP+FP+TN+FN} \tag{13}$$

Equation (13) defines accuracy as the ratio of correctly categorised data to all categorised data. The variables TP, TN, FN, and FP represent true positives, true negatives, false negatives, and false positives, respectively. True positives represent the positive class that is predicted to be positive; true negatives represent the negative class that is predicted to be negative; false negatives represent the positive class that is predicted to be negative; and false positives represent the negative class that is predicted to be positive.

Accuracy is the ratio of the number of correctly categorized samples to the total number of samples. A higher accuracy rate means that a higher percentage of people with good and bad credit are successfully identified.

The mathematical definition of recall is as follows

$$Recall=\frac{TP}{TP+FN} \tag{14}$$

Equation (14) defines recall as the ratio of correctly predicted positive categories to all the actual positive categories in the data.

A 100% recall rate signifies that no benign credit applicant is rejected erroneously. For example, commercial banks generate the majority of their profits from the deposit credit spread. Consequently, wrongly rejecting a benign loan application results in loss of profit. In situations where a bank seeks to expand its lending business, the cost of wrongly rejecting a conscientious loan application is greater than that of passing a non-performing load application. Therefore, at this juncture, it can be concluded that an elevated recall rate engenders a more beneficial model.

The mathematical definition of precision is as follows

$$Precision=\frac{TP}{TP+FP} \tag{15}$$

TABLE I
EXPERIMENTAL RESULTS FOR THE ACCURACY

Dataset	LDA	LR	SVM
German credit dataset (Original data)	75.1%	75.6%	75.1%
bin=3	76.3%	76.5%	76.0%
bin=4	76.9%	77.3%	76.5%
bin=5	75.5%	75.9%	75.7%
bin=6	76.0%	76.2%	76.3%
bin=7	75.9%	75.9%	76.4%
bin=8	76.8%	76.8%	77.2%
bin=9	76.4%	76.5%	76.7%
bin=10	76.9%	76.6%	76.3%
bin=11	75.6%	76.1%	76.1%
bin=12	76.9%	76.1%	76.4%
bin=13	76.5%	77.4%	76.8%
bin=14	76.2%	76.3%	76.3%
bin=15	76.5%	76.8%	77.2%
bin=16	76.3%	76.6%	76.1%
Average	76.3%	76.5%	76.4%

Equation (15) defines precision as the ratio of correctly predicted positive categories to all predicted positive categories in the data.

Precision represents the percentage of applicants with genuinely good credit among those rated as having good credit. A higher precision value indicates a lower percentage of non-performing loans formed from all approved applications, which in turn implies a corresponding reduction in the incidence of bad debts in the future.

The mathematical definition of F1-value is as follows

$$F_1 = 2\frac{Precision \times Recall}{Precision + Recall} \quad (16)$$

Equation (16) defines the F1-value as a measure that balances the recall and precision to illustrate the classification effect.

F1-value is a measure of both precision and recall, with higher values indicating superior classification performance.

Higher values of the four metrics (accuracy, recall, precision, and F1-value) indicate better classification performance.

The experiments involved transforming 1000 personal credit data from the German credit dataset into a $4 \times 5 \times 1000$ CELL matrix. The data was then classified using three algorithms: LDA, LR, and SVM. The values of bin ranged from 3 to 16.

The experiment utilised a 10-fold cross-validation. The German credit dataset contains 1000 individual credit data, which were evenly divided into 10 parts, each with 100 individual credit data. Ten experiments were conducted on the 10 individual credit data respectively. Each time, one of the 10 parts of the data was used as the test data in turn, and the remaining 9 parts of the data were used as the training data. Subsequently, the experiment involved the utilisation of ten copies of data, which were employed in succession as test data. The average of the classification results were then calculated, thereby providing an estimation of the utility of the method.

Table I shows the experimental results of the accuracy.

Table I summarizes the comparison results of the average classification accuracy of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table I, we can see that the average classification accuracies of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 1.2%, 0.9%, and 1% higher than that of the German credit dataset, respectively.

TABLE II
EXPERIMENTAL RESULTS FOR THE RECALL

Dataset	LDA	LR	SVM
German credit dataset (Original data)	87.4%	88.0%	88.7%
bin=3	88.3%	88.3%	89.9%
bin=4	88.6%	88.6%	89.1%
bin=5	88.4%	88.6%	89.6%
bin=6	88.0%	88.4%	89.6%
bin=7	88.0%	88.4%	90.1%
bin=8	88.3%	88.4%	89.9%
bin=9	88.6%	88.7%	89.7%
bin=10	88.7%	88.9%	90.1%
bin=11	88.7%	88.1%	89.4%
bin=12	88.6%	88.2%	89.3%
bin=13	88.0%	88.9%	90.3%
bin=14	88.3%	88.3%	90.0%
bin=15	88.0%	88.4%	89.9%
bin=16	88.0%	88.7%	89.9%
Average	88.2%	88.5%	89.8%

Table II shows the experimental results of the recall.

Table II summarizes the comparison results of the average classification recall of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table II, we can see that the average classification recalls of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 0.8%, 0.5%, and 1% higher than that of the German credit dataset, respectively.

TABLE III
EXPERIMENTAL RESULTS FOR THE PRECISION

Dataset	LDA	LR	SVM
German credit dataset (Original data)	79.2%	79.4%	78.5%
bin=3	79.9%	80.2%	78.8%
bin=4	80.4%	80.8%	79.7%
bin=5	79.4%	79.5%	78.7%
bin=6	79.8%	79.8%	79.3%
bin=7	79.7%	79.5%	79.1%
bin=8	80.5%	80.4%	80.0%
bin=9	79.9%	79.9%	79.6%
bin=10	80.3%	79.9%	79.0%
bin=11	79.5%	80.0%	79.1%
bin=12	80.4%	80.1%	79.5%
bin=13	80.3%	80.8%	79.4%
bin=14	79.8%	80.3%	79.0%
bin=15	80.3%	80.4%	80.0%
bin=16	80.1%	80.0%	78.9%
Average	80.0%	80.1%	79.3%

Table III shows the experimental results of the precision.

Table III summarizes the comparison results of the average classification precision of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table III, we can see that the average

TABLE IV
EXPERIMENTAL RESULTS FOR THE F1-VALUE

Dataset	LDA	LR	SVM
German credit dataset (Original data)	83.1%	83.5%	83.3%
bin=3	83.9%	84.0%	84.0%
bin=4	84.3%	84.5%	84.2%
bin=5	83.5%	83.7%	83.8%
bin=6	83.7%	83.9%	84.1%
bin=7	83.6%	83.7%	84.2%
bin=8	84.2%	84.2%	84.7%
bin=9	84.0%	84.1%	84.4%
bin=10	84.3%	84.2%	84.2%
bin=11	83.4%	83.7%	84.0%
bin=12	84.3%	83.7%	84.1%
bin=13	84.0%	84.6%	84.5%
bin=14	83.9%	83.8%	84.2%
bin=15	84.0%	84.2%	84.7%
bin=16	83.9%	84.1%	84.0%
Average	84.0%	84.0%	84.2%

classification precisions of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 0.8%, 0.7%, and 0.8% higher than that of the German credit dataset, respectively.

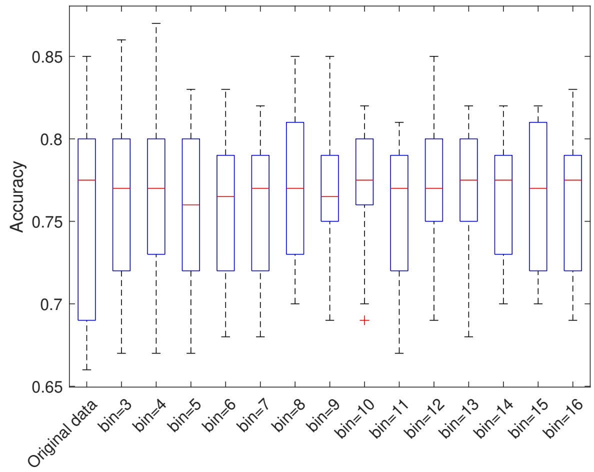
Table IV shows the experimental results of the F1-value.

Table IV summarizes the comparison results of the average classification F1-value of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table IV, we can see that the average classification F1-values of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 0.8%, 0.5%, and 0.9% higher than that of the German credit dataset, respectively.

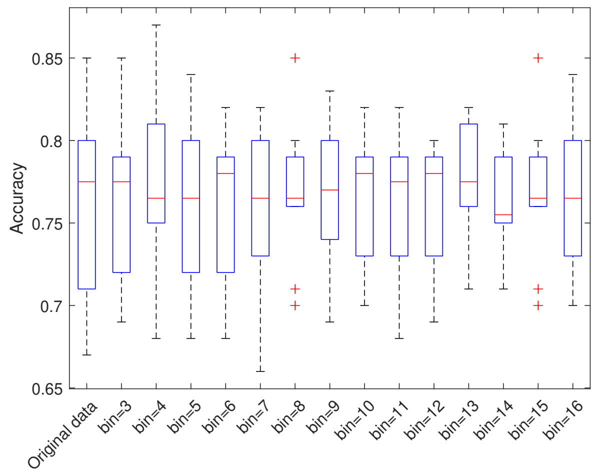
To summarise, the German credit dataset based on HOG yields higher accuracies, recalls, precisions, and F1-values compared to the German credit dataset when employing LDA, LR, and SVM algorithms.

The 10-fold cross-validation allows for the derivation of average values from the data presented in Tables I to IV. The construction of the corresponding box plots for the results of the 10 experiments provides a means of more accurately reflecting the performance of the method under different algorithms, as well as its stability. This includes the ability to identify outliers, assess the distribution of data, and determine whether the date is centralized or decentralized, as illustrated in Figure 2, 3, 4, and 5.

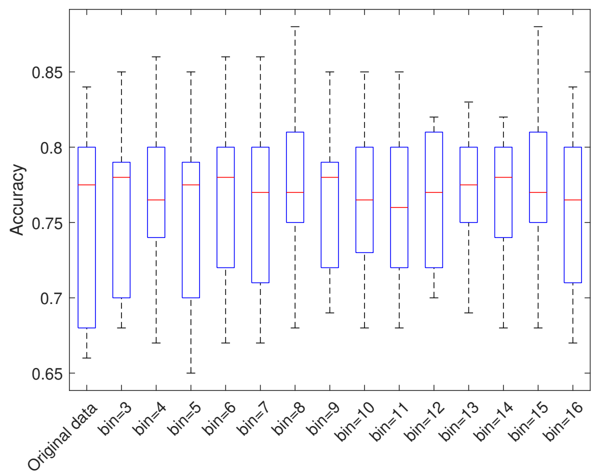
Figure 2 shows box plots of accuracy based on LDA, LR, and SVM algorithms.



(a) Box plot of accuracy based on LDA algorithm.



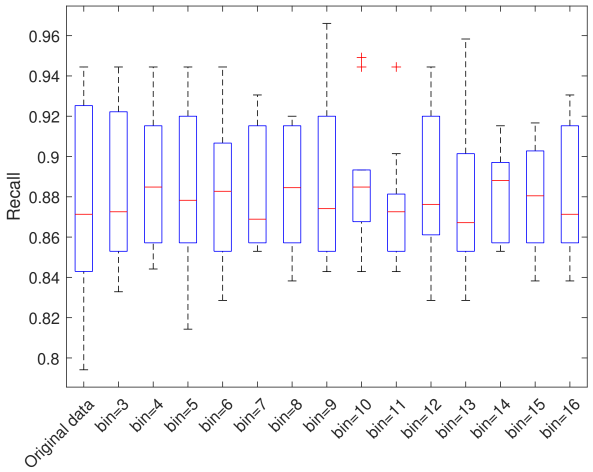
(b) Box plot of accuracy based on LR algorithm.



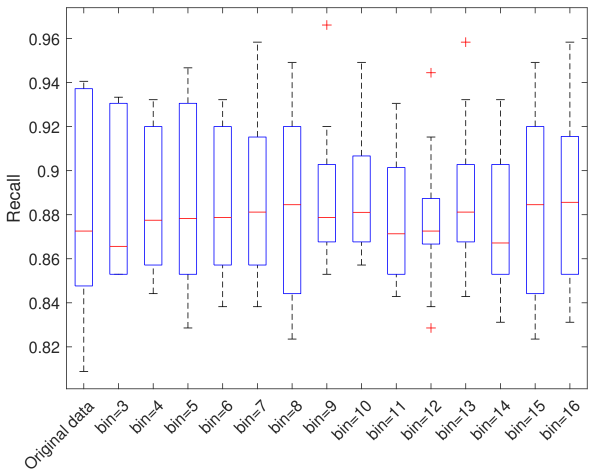
(c) Box plot of accuracy based on SVM algorithm.

Fig. 2. Box plots of accuracy based on LDA, LR, and SVM algorithms.

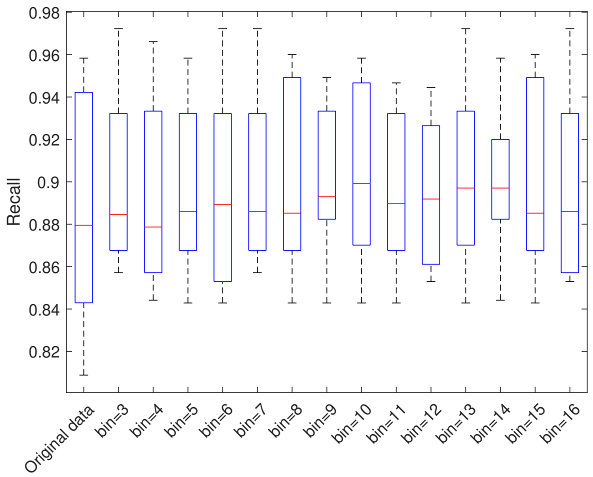
As illustrated in Figure 2, the application of LDA, LR, and SVM algorithms in conjunction with the German credit dataset based on HOG exhibits a smaller Interquartile Range (IQR) in terms of accuracy when compared to the German credit dataset. In Figure 2(a), one outlier occurs when bin = 8. However, the outliers fall within the range of the maximum and minimum values of the German credit dataset box plot, and the IQR of the box plot for bin = 8 is considerably smaller than that of the box plot for the German credit dataset. Similarly, in Fig. 2(b), there are six outliers when bin=8 and 15. The outliers also fall within the range of the maximum and minimum values of the German credit dataset box plot, and the IQR of the box plot for bin = 8 and 15 are also considerably smaller than that of the box plot for the German credit dataset. The results of the experimental comparisons demonstrate that the German credit dataset based on HOG has superior stability in terms of accuracy



(a) Box plot of recall based on LDA algorithm.



(b) Box plot of recall based on LR algorithm.



(c) Box plot of recall based on SVM algorithm.

Fig. 3. Box plots of recall based on LDA, LR, and SVM algorithms.

when compared to the German credit dataset.

Figure 3 shows the box plots of recall based on LDA, LR, and SVM algorithms.

As illustrated in Figure 3, the application of LDA, LR, and SVM algorithms in conjunction with the German credit dataset based on HOG exhibits a smaller IQR in terms of recall when compared to the German credit dataset. In Figure 3(a), three outliers occur when bin = 10 and 11. However, the outliers fall within the range of the maximum and minimum values of the German credit dataset box plot, and the IQR of the box plots for bin = 10 and 11 are considerably smaller than those of the box plot for the German credit dataset. In Fig. 3(b), there are four outliers when bin=9, 12, and 13. One outlier falls within the range of the maximum and minimum values of the German credit dataset box plot. The other three outliers exceed the range of the maximum values of the German credit dataset boxplot, and exhibit a rightward skew because of the considerably smaller IQR in comparison to the German credit dataset. The results of the experimental comparisons demonstrate that the German credit dataset based on HOG has superior stability in terms of recall when compared to the German credit dataset.

Figure 4 shows the box plots of precision based on LDA, LR, and SVM algorithms.

As illustrated in Figure 4, the application of LDA, LR, and SVM algorithms in conjunction with the German credit dataset based on HOG exhibits a smaller IQR in terms of precision when compared to the German credit dataset. Moreover, there are no outliers in Figure 4. The results of the experimental comparisons demonstrate that the German credit dataset based on HOG has superior stability in terms of precision when compared to the German credit dataset.

Figure 5 shows the box plots of F1-value based on LDA, LR, and SVM algorithms.

As illustrated in Figure 5, the application of LDA, LR, and SVM algorithms in conjunction with the German credit dataset based on HOG exhibits a smaller IQR in terms of F1-value when compared to the German credit dataset. In Figure 5(a), one outlier occurs when bin = 9. However, the outlier falls within the range of maximum and minimum values of the German credit dataset boxplot, and the IQR of the box plot for bin = 9 is considerably smaller than that of the box plot for the German credit dataset. The results of the experimental comparisons demonstrate that the German credit dataset based on HOG has superior stability in terms of F1-value when compared to the German credit dataset.

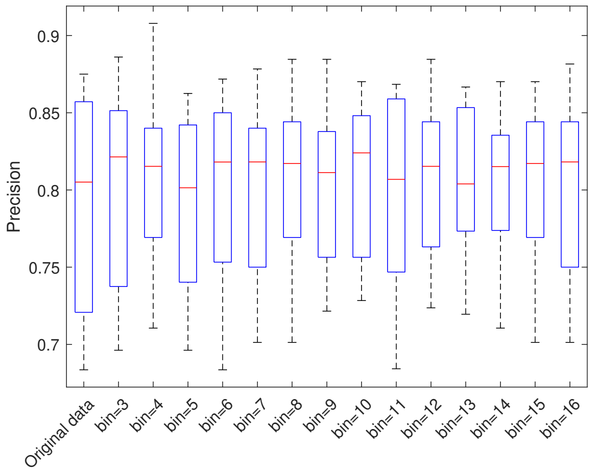
In addition to box plots, standard deviation also can be utilised to ascertain the stability of the results of the 10 experiments.

Table V shows the standard deviation results of the accuracy.

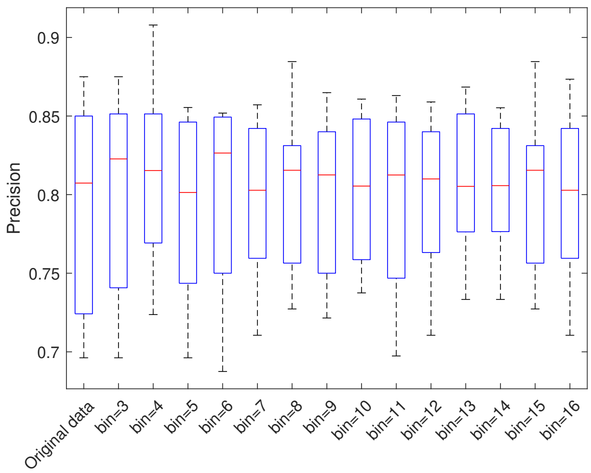
Table V summarizes the comparison results of the standard deviation for accuracy of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table V, we can see that the average standard deviations for accuracy of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 0.0127, 0.0111, and 0.0054 higher than that of the German credit dataset, respectively.

Table VI shows the standard deviation results of the recall.

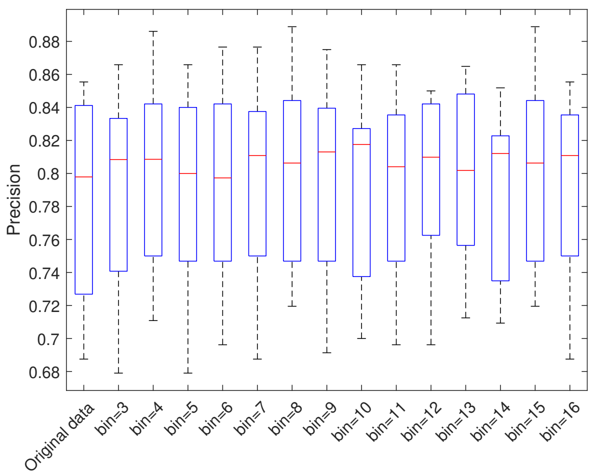
Table VI summarizes the comparison results of the standard deviation for recall of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table II, we can see that the average standard deviations for recall of the German credit dataset based on HOG using LDA, LR, and SVM



(a) Box plot of precision based on LDA algorithm.

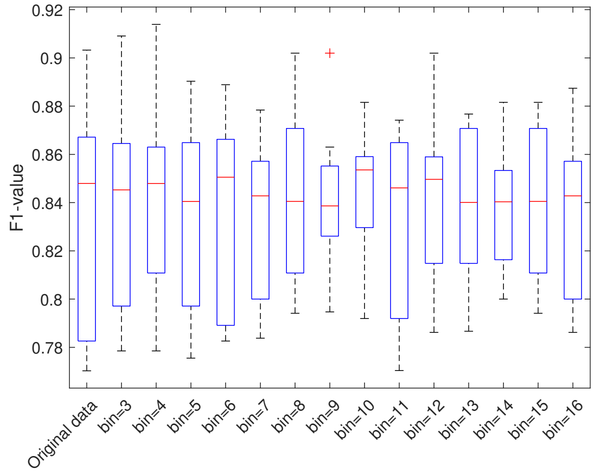


(b) Box plot of precision based on LR algorithm.

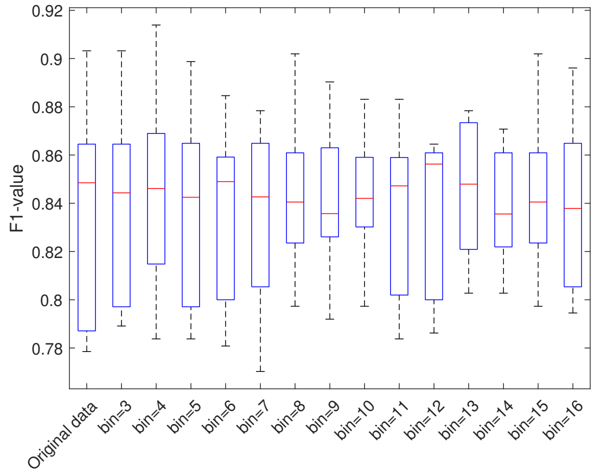


(c) Box plot of precision based on SVM algorithm.

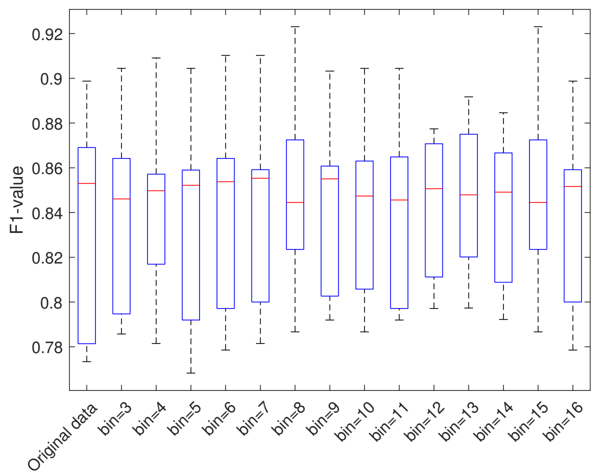
Fig. 4. Box plots of precision based on LDA, LR, and SVM algorithms.



(a) Box plot of F1-value based on LDA algorithm.



(b) Box plot of F1-value based on LR algorithm.



(c) Box plot of F1-value based on SVM algorithm.

Fig. 5. Box plots of F1-value based on LDA, LR, and SVM algorithms.

algorithms are 0.0137, 0.0068, and 0.0075 higher than that of the German credit dataset, respectively.

Table VII shows the standard deviation results of the precision.

Table VII summarizes the comparison results of the standard deviation for precision of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table I, we can see that the average standard deviations for precision of the German credit dataset based on HOG using LDA, LR, and SVM

TABLE V
STANDARD DEVIATION FOR THE ACCURACY

Dataset	LDA	LR	SVM
German credit dataset (Original data)	0.0587	0.0541	0.0558
bin=3	0.0560	0.0490	0.0501
bin=4	0.0530	0.0514	0.0520
bin=5	0.0516	0.0513	0.0547
bin=6	0.0500	0.0469	0.0550
bin=7	0.0439	0.0455	0.0541
bin=8	0.0442	0.0407	0.0545
bin=9	0.0439	0.0437	0.0478
bin=10	0.0411	0.0375	0.0504
bin=11	0.0504	0.0483	0.0505
bin=12	0.0468	0.0394	0.0418
bin=13	0.0427	0.0358	0.0442
bin=14	0.0352	0.0287	0.0450
bin=15	0.0393	0.0407	0.0545
bin=16	0.0452	0.0434	0.0507
Average	0.0460	0.0430	0.0504

TABLE VI
STANDARD DEVIATION RESULTS FOR THE RECALL

Dataset	LDA	LR	SVM
German credit dataset (Original data)	0.0459	0.0404	0.0456
bin=3	0.0380	0.0336	0.0387
bin=4	0.0308	0.0319	0.0401
bin=5	0.0383	0.0383	0.0384
bin=6	0.0341	0.0323	0.0411
bin=7	0.0274	0.0347	0.0384
bin=8	0.0299	0.0398	0.0423
bin=9	0.0376	0.0327	0.0341
bin=10	0.0332	0.0270	0.0396
bin=11	0.0276	0.0275	0.0346
bin=12	0.0379	0.0320	0.0317
bin=13	0.0379	0.0322	0.0399
bin=14	0.0201	0.0312	0.0308
bin=15	0.0276	0.0398	0.0423
bin=16	0.0299	0.0376	0.0411
Average	0.0322	0.0336	0.0381

TABLE VII
STANDARD DEVIATION RESULTS FOR THE PRECISION

Dataset	LDA	LR	SVM
German credit dataset (Original data)	0.0622	0.0587	0.0570
bin=3	0.0613	0.0585	0.0562
bin=4	0.0584	0.0565	0.0565
bin=5	0.0565	0.0544	0.0553
bin=6	0.0564	0.0569	0.0569
bin=7	0.0552	0.0506	0.0553
bin=8	0.0528	0.0459	0.0532
bin=9	0.0528	0.0509	0.0557
bin=10	0.0488	0.0443	0.0552
bin=11	0.0606	0.0556	0.0564
bin=12	0.0503	0.0493	0.0513
bin=13	0.0502	0.0448	0.0517
bin=14	0.0474	0.0398	0.0491
bin=15	0.0508	0.0459	0.0532
bin=16	0.0555	0.0490	0.0541
Average	0.0541	0.0502	0.0543

TABLE VIII
STANDARD DEVIATION RESULTS FOR THE F1-VALUE

Dataset	LDA	LR	SVM
German credit dataset (Original data)	0.0432	0.0391	0.0412
bin=3	0.0393	0.0350	0.0374
bin=4	0.0373	0.0356	0.0354
bin=5	0.0369	0.0374	0.0403
bin=6	0.0367	0.0335	0.0400
bin=7	0.0312	0.0333	0.0384
bin=8	0.0329	0.0300	0.0395
bin=9	0.0300	0.0296	0.0341
bin=10	0.0281	0.0249	0.0344
bin=11	0.0376	0.0355	0.0351
bin=12	0.0343	0.0301	0.0304
bin=13	0.0312	0.0273	0.0315
bin=14	0.0254	0.0212	0.0304
bin=15	0.0295	0.0300	0.0395
bin=16	0.0330	0.0328	0.0367
Average	0.0331	0.0312	0.0360

algorithms are 0.0082, 0.0085, and 0.0027 higher than that of the German credit dataset, respectively.

Table VIII shows the standard deviation results of the F1-value.

Table VIII summarizes the comparison results of the standard deviation for F1-value of the German credit dataset and the German credit dataset based on HOG using LDA, LR, and SVM algorithms. From Table IV, we can see that the average standard deviations for F1-value of the German credit dataset based on HOG using LDA, LR, and SVM algorithms are 0.8%, 0.5%, and 0.9% higher than that of the German credit dataset, respectively.

In conclusion, the experimental results indicate that the accuracy, recall, precision, and F1-value of the German credit dataset based on HOG exhibit superior stability to the German credit dataset when utilising LDA, LR, and SVM algorithms.

IV. Conclusion

In this study, a method based on HOG was proposed for personal credit scoring. We employed the gradient relationship between disparate features of personal credit data and introduced novel features derived from the original features of personal credit data to construct a set of personal credit datasets based on HOG. The experimental results demonstrate that the German credit dataset based on HOG achieves a better performance compared to the German credit dataset in both terms of classification and stability.

In future work, we will not only consider the German credit dataset, but also different datasets such as the Australian credit dataset, the Bene 1 dataset, the European bank dataset, the Hong Kong bank dataset, and other different datasets for experiments, so as to study more fully the impacts of diverse cultural backgrounds and credit scoring factors on individual credit scores at greater depths.

References

[1] D. West, "Neural network credit scoring models," *Computers and Operations Research*, vol. 27, no. 11-12, pp. 1131-1152, 2000.

[2] T. S. Lee, C. C. Chiu, C. J. Lu, and I. F. Chen, "Credit scoring using the hybrid neural discriminant technique," *Expert Systems with Applications*, vol. 23, no. 3, pp. 245-254, 2002.
[3] B. Baesens, R. Setiono, C. Mues, and J. VanthienenT, "Using neural network rule extraction and decision tables for credit-risk evaluation," *Management Science*, vol. 49, no. 3, pp. 312-329, 2003.
[4] C. S. Ong, J. J. Huang, and G. H. Tzengb, "Building Credit Scoring Models Using Genetic Programming," *Expert Systems with Applications*, vol. 29, no. 1, pp. 41-47, 2005.
[5] C. L. Huang, M. C. Chen, and C. J. Wang, "Credit Scoring with a Data Mining Approach Based on Support Vector Machines," *Expert Systems with Applications*, vol. 33, no. 4, pp. 847-856, 2007.
[6] M. Mavri, V. Angelis, G. Ioannou, E. Gaki, and I. Koufodontis, "A two-stage dynamic credit scoring model, based on customers' profile and time horizon," *Journal of Financial Services Marketing*, vol. 13, no. 1, pp. 17-27, 2008.
[7] L. Yu, S. Wang, and K. K. Lai, "Credit risk assessment with a multistage neural network ensemble learning approach," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1434-1444, 2008.
[8] A. C. Antonakis, and M. E. Sfakianakis, "Assessing naive Bayes as a method for screening credit applicants," *Journal of Applied Statistics*, vol. 36, no. 5, pp. 537-545, 2009.
[9] Y. Zhao, Y. Zhao, and I. Song, "Predicting New Customers' Risk Type in the Credit Card Market," *Journal of Marketing Research*, vol. 46, no. 4, pp. 506-517, 2009.
[10] L. Zhou, K. K. Lai, and J. Yen, "Credit Scoring Models with AUC Maximization Based on Weighted SVM," *International Journal of Information Technology and Decision Making*, vol. 8, no. 44, pp. 677-696, 2009.
[11] C. L. Chuang, and S. T. Huang, "A hybrid neural network approach for credit scoring," *Expert systems: The international journal of knowledge engineering*, vol. 28, no. 2, pp. 185-196, 2011.
[12] Y. Wei, P. Yildirim, C. Van Den Bulte, and C. Dellarocas, "Credit Scoring with Social Network Data," *Marketing Science*, vol. 35, no. 2, pp. 224-258, 2016.
[13] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, "Advances in Social Media Research: Past, Present and Future," *Information Systems Frontiers*, vol. 20, no. 3, pp. 531-558, 2018.
[14] X. Fu, T. Ouyang, J. Chen, and X. Luo, "Listening to the investors: A novel framework for online lending default prediction using deep learning neural networks," *Information Processing and Management*, vol. 57, no. 4, Article ID102236, 13 pages, 2020.
[15] C. Zhang, Z. Yang, X. He, and L. Deng, "Multimodal intelligence: Representation learning, information fusion, and applications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 3, pp. 478-493, 2020.
[16] Y. Li, and G. Wen, "Research and Practice of Financial Credit Risk Management Based on Federated Learning," *Engineering Letters*, vol. 31, no. 1, pp. 271-278, 2023.
[17] K. Zhou, C. Zhang, Y. Yu, S. Cong, and X. Yue, "Improving SMOTE Technology for Credit Card Fraud Detection Category Imbalance Issues," *Engineering Letters*, vol. 31, no. 4, pp. 1780-1785, 2023.
[18] F. M. Talaat, A. Aljadani, M. Badawy, and M. Elhosseini, "Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction," *Journal of Financial Stability*, vol. 9, no. 8, pp. 4847-4865, 2024.
[19] S. J. Bennehalli, S. Vakkund, A. Hegde and B. Bhowmik, "Navigating Data Imbalances in Credit Risk Management: A One-Sided Selection Approach," in *Proceedings of Conference on Control Instrumentation System 2024*, pp. 1-6.
[20] J. Chen, G. Tian, and J. Wang, "Application of Big Data Technology in Internet Financial Risk Control," *IAENG International Journal of Computer Science*, vol. 51, no. 8, pp. 1155-1162, 2024.
[21] A. M. Idrees, N. S. Elhusseny, and S. Ouf , "Credit Card Fraud Detection Model-based Machine Learning Algorithms," *IAENG International Journal of Computer Science* , vol. 51, no. 10, pp. 1649-1662, 2024.
[22] T. Y. Chiang, Z. Luo, H. Li, C. Zhang, W. Xie, and L. Zhou, "Inventory Systems with Present Value and Credit Period," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 11, pp. 2279-2289, 2024.
[23] Z. Meng, X. Xie, Y. Xie. and J. Sun, "Federated Learning for Short-term Load Forecasting," *IAENG International Journal of Computer Science*, vol. 52, no. 1, pp. 52-58, 2025.
[24] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of Conference on Computer Vision and Pattern Recognition 2005*, pp. 886-893.