

ISSN 2185-3118

INTERNATIONAL JOURNAL OF INTELLIGENT ENGINEERING AND SYSTEMS

Volume 1 Issue 1 March 2008



Intelligent Networks and Systems Society



The Intelligent Networks
And Systems Society

Publications

Editorial Staff

[Publications
menu](#)

**International Journal of Intelligent Engineering and
Systems**

indexed by,

| SCOPUS

| Scimago

| Crossref

| EBSCOhost

| Ulrich's

| OAJI

UGC Approved Journal (Journal No. 22983)

Quarterly e-Journal (From 2008 - 2016)

Bi-Monthly e-Journal (Since 2017)

ISSN: 2185-3118

Advisory Board

Hirofumi Sasaki, Professor Emeritus of Tokai University, Japan

Hassan Bevrani, University of Kurdistan, Iran

Editor-in-Chief

Kei Eguchi, Fukuoka Institute of Technology, Japan

Associate Editors-in-Chief

Amphawan Julsereewong, King Mongkut's Institute of Technology Ladkrabang, Thailand

Sawai Pongswatd, King Mongkut's Institute of Technology Ladkrabang, Thailand

Editorial Board

Nam-Soo Kim, Cheongju University, Republic of Korea

B. T. P. Madhav, K L University, India

Farzin Asadi, Kocaeli University, Turkey

Win Pa Pa, University Computer Studies Yangon, Myanmar

Prasit Julsereewong, King Mongkut's Institute of Technology Ladkrabang, Thailand

Worapong Tangsirat, King Mongkut's Institute of Technology Ladkrabang, Thailand

Prapart Ukakimaparn, King Mongkut's Institute of Technology Ladkrabang, Thailand

Sawatsakorn Chaiyasoonthorn, Ramkhamhaeng University, Thailand

Krit Smerpitak, King Mongkut's Institute of Technology Ladkrabang, Thailand

Apinai Rerkratn, King Mongkut's Institute of Technology Ladkrabang, Thailand

Pittaya Pannil, King Mongkut's Institute of Technology Ladkrabang, Thailand

Anucha Kaewpoonsuk, Naresuan University, Thailand

Teerawat Thepmanee, King Mongkut's Institute of Technology Ladkrabang, Thailand

Chirdpong Deelertpaiboon, King Mongkut's University of Technology North Bangkok, Thailand

Natee Thong-un, King Mongkut's University of Technology North Bangkok, Thailand

Sart Kummool, King Mongkut's Institute of Technology Ladkrabang, Thailand

Ichirou Oota, National Institute of Technology Kumamoto College, Japan

Takaaki Ishibashi, National Institute of Technology Kumamoto College, Japan

Kou Yamada, Gunma University, Japan

Kuniaki Fujimoto, Tokai University, Japan

Mitsutoshi Yahara, Tokai University Fukuoka Junior College, Japan

Shuji Kurebayashi, Shizuoka University, Japan

Wanglok Do, Fukuoka Institute of Technology, Japan

Riyanarto Sarno, Institut Teknologi Sepuluh Nopember, Indonesia

Board member responsibilities

* Editorial board members must obey the direction provided by the Editors-in-Chief of the journal.

* Editorial board members must judge every submission equally and transparently.

- * Editorial board members should not reveal any information regarding the author, reviewer or the paper information to anyone.
- * Editorial board members are responsible for timely publishing the accepted papers.
- * Editorial board members may attend the Editorial Board meetings occasionally scheduled by the Editorial Office for taking part in the discussion for the improvement of the Journal.
- * Editors should not have any conflict of interest with any assignment.
- * Editors should be responsible for a fast and transparent peer review process.
- * Editors should keep the policy of fast and effective peer review.
- * Editors should be responsible for any query regarding reconsideration of editorial decisions and should provide the decision quickly and clearly with proper reasons.
- * Editors should be responsible for selecting potential reviewers considering their expertise in particular subject areas and keeps monitoring the review process.
- * Editors will be responsible to convey the expectations of the journal to the reviewers with the review scope, quality and timeliness for an effective, fair and constructive review for the assigned submission.
- * Editors should keep in mind the time required for reviewing papers before sending any reminder to the reviewer so that assigned reviewer should get appropriate time he or she requires.
- * Editors-in-Chief can take the final decision for any publication oriented issue.



Modified Focal Loss in Imbalanced XGBoost for Credit Card Fraud Detection

Dedy Trisanto^{1*}Nofita Rismawati²Muhamad Femy Mulya³Felix Indra Kurniadi³¹ Politeknik Sekolah Tinggi Manajemen Industri Jakarta, Indonesia² Indraprasta Persatuan Guru Republik Indonesia University, Indonesia³ Tanri Abeng University, Indonesia* Corresponding author's Email: dedy.trisanto@stmi.ac.id

Abstract: The development of credit card use in Indonesia has not been matched by the security provided by credit card service providers. This resulted in significant losses both in terms of banking and customers. The difficulty in finding the characteristics of credit card fraud is one of the biggest challenges. Currently, many are developing machine learning models that can identify credit card fraud to help banks. Unfortunately, the model created is mostly biased towards the class which has more dominant data. This problem is caused by the imbalance of the data on the available dataset. In the previous research, Wang et al found the implementation of Focal loss in XGBoost improve the precision, recall of imbalanced data. However according to Qin et al, the parameter loss from Focal loss have poor judgement in several case of imbalanced data and to handle this weakness, they proposed Weighted – Cross Entropy Loss (W-CEL) loss in Focal loss. According to previous research, we propose a Modified Focal Loss method for Imbalanced XGBoost by entering another parameter from W-CEL loss to Focal Loss to improve the ability of Focal Loss. Focal loss itself is a method that is often used to give weight to classes that are often misinterpreted, so that with the use of imbalance parameters (ϕ) from W-CEL loss is expected to improve the weighted value of the Focal Loss. We tested our proposed method in credit card fraud dataset from Université Libre de Bruxelles (ULB) machine learning group. Our proposed method produced 100 % accuracy, 0.97 precision, 0.56 recall and 0.72 MCC score in scenario 1 with extreme imbalanced data, in scenario 2 with the mild imbalanced data, the result delivered 99% accuracy, 0.88 precision, 0.87 recall and 0.89 MCC and in scenario 3 with the medium imbalance data, the result delivered 100% accuracy, 0.97 precision score, 0.72 recall score and 0.83 MCC score. The results obtained in this study proved that the proposed method makes machine learning models valid and unbiased, especially in mild-imbalanced data. However, many improvements still need to be made to medium and extreme imbalanced data.

Keywords: Fraud detection, XGBoost, Focal loss, Weighted cross entropy loss (W-CEL), Imbalanced XGBoost.

1. Introduction

The pandemic period that has occurred since 2020 until recently has changed the paradigm of transactions in Indonesia, especially from conventional transactions to online ones. This has also begun to change the way payments are made, where many people pay using online money or credit cards. In fact, based on data provided by Bank Indonesia (BI), there is an increase in the number of credit cards in circulation with an increase of 2.67% in 2020 [1].

Unfortunately, the development of this increase in users has not been matched by security for credit card users in Indonesia. In fact, based on articles published by CNBC in Forter's Analysis, there are 5 worst countries in driving away e-commerce fraud and Indonesia is ranked first in that category. It can be seen that there is a need for an early detection to ward off possible losses incurred in settling credit card fraud [2].

Several studies have tried to overcome problems in problem detection in credit card fraud such as [3] using traditional machine learning methods such as Neural Network, Support Vector Machine, K-Nearest Neighbor, and Logistic Regression. Some use the

ensemble method to improve the accuracy of the proposed machine learning models such as that of [4-7].

The accuracy results provided by each method that have been proposed in previous studies have even reached more than 90% of the accuracy level. Unfortunately, when the results are seen in the precision, recall and Matthew Correlation Coefficient (MCC) values, it is found that the values of the three evaluations are very inversely proportional to their accuracy values. It can be seen in Table 1 in section 4 that the values of the three evaluations are lower than the available accuracy values. This is due to the problematic data imbalance in the dataset. The imbalance case in this dataset gives bias to one of the classes that has more dominant data so that the results of the model created cannot be considered valid.

Several studies suggest some solutions in the pre-process such as [8, 9] proposing using the oversampling method, and [9] proposing using the under-sampling method. In addition to the sampling method, there are several researchers who propose a combination of outlier detection and sampling, such as [10]. The approach taken by previous researchers is a good one, especially in terms of imbalance. However, doing the under-sampling method will reduce important data which sometimes can become main information or important data that can be analyzed more deeply. While the oversampling method will provide data that is similar to the existing data, the data is only used as augmentation data so that the data does not become an imbalance.

This problem is what makes author interested in increasing the value of the three evaluations such as precision, recall and, MCC without doing pre-processing such as sampling or hybrid sampling and outlier detection.

The XGBoost method is one of the most frequently used methods for solving fraud detection. Although the performance of the XGBoost algorithm is very good in the classification and regression process, sometimes it is difficult to get maximum results, especially in the case of imbalance.

Several researchers have tried to improve the XGBoost method to make it more robust against imbalance data, one of which is the method proposed by Wang et.al. [11] using weighted (cross-entropy) and focal losses in the boosting machine. The choice of weighted cross entropy as the main method was chosen because it is one of the good and easy algorithms in the cost sensitive method of data imbalance and focal loss. This method was chosen

because it reduces the importance of well-classified data points which are very useful for imbalance data.

Another study proposed by Priscilla et al. which proposes a tuning optimizer on the hyperparameter using Randomized Search [12]. These two studies provide an overview of the usefulness of hyperparameters and loss in the application of XGBoost in the Imbalance dataset. Focal loss application requires hyper-parameters γ to determine the best value of the selected gamma. Besides that, based on Qin et al [13], the value of imbalanced parameter ϕ which is at the Weighted-Cross Entropy Loss (W-CEL) can be implemented Focal loss to help overcome the shortcomings of imbalance parameter γ in Focal loss.

The goals and contributions given by the author in this scientific article are:

- To use imbalanced parameter ϕ from W-loss in Focal loss
- To use tuning-hyperparameter on γ value.
- To evaluate the model of Modified Focal Loss on XGBoost Imbalance using evaluation metrics: precision, recall, and MCC.

This article consists of the first section which discusses the background of this research, Section 2 which discusses the data used, Section 3 which discusses the research methodology carried out, Section 4 which discusses the experiment and results, and Section 5 which discusses the conclusions of research conducted.

2. Data

Information about credit cards is one of the most important consumer information and cannot be shared carelessly. This has resulted in us using the publicly available data for this study.

The dataset used in this experiment is the credit card fraud dataset provided by the Université Libre de Bruxelles (ULB) machine learning group. This dataset contains the use of credit cards in Europe for two days where the sampling was carried out in September 2013. The data obtained from this sampling were 284'807, where transactions with 492 data proved to be fraud representing 0.172% of the data, and the rest is a normal transaction. It can be seen from this description that this data is data that has extreme imbalance[14]. The Dataset was downloaded in datahub.io¹

The dataset itself consists of 28 features that are principal components from the results of using the Principal Component Analysis method on raw data. Principal analysis method is used to protect

¹ <https://datahub.io/machine-learning/creditcard>

consumers so that data secrets are not exposed. Another feature is "time" and "amount", so there are a total of 30 features [15].

3. Research methodology

The research stages in this article are divided into two main parts, which are data pre-processing and detection using the Modified Imbalanced XGBoost method.

3.1 Pre-processing

Pre-processing is one of the important steps that need to be taken in order to make data more valid and unbiased. The initial stage in pre-processing is normalizing the data amount and making some unused features such as "time". The normalization process is carried out using the min-max scaler method, the formula for the min-max scaler is [16]:

$$d' = \frac{d - \min(d)}{\max(d) - \min(d)} \quad (1)$$

Where d' is the normalized feature data value, d is the raw feature data value, $\min(d)$ is the minimum value of all data, $\max(d)$ is the maximum value of all data.

After the 'amount' feature is normalized, the next step is to find the most relevant features to use. Thus, to find the uniqueness of each feature, we use the Correlation Coefficient to find the best features of the existing features [17].

$$r = \frac{\sum(F_i^1 - \bar{F}^1)(F_i^2 - \bar{F}^2)}{\sqrt{\sum(F_i^1 - \bar{F}^1)^2 \sum(F_i^2 - \bar{F}^2)^2}} \quad (2)$$

Where F^1, F^2 is feature 1 and feature 2, and \bar{F}^1, \bar{F}^2 is the average of feature 1 and feature 2

After completing the pre-processing of data, the next step is to create a model using the XGBoost Imbalance method.

3.2 XGBoost

XGBoost is a machine learning method based on tree boosting introduced by Chen et al [18]. XGBoost is a popular algorithm that adopts an additive learning scheme with second-order approximation, first order derivative which is called gradient and second order derivative which is called hessian of loss function.

$$L^{(t)} = \sum_{i=1}^n l(y_i, p_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (3)$$

The Eq. (3) is a regularized objective where $l(\dots, \dots)$ is a loss between target y_i , and prediction (p_i). Input is symbolized with (x_i). The complexity of the model is symbolized with $\Omega(\dots)$, m is the number of data and n is the number of features. We used additive manner for fitted tree $p_i^{(t)} = p_i^{(t-1)} + f_t(x_i)$ where t is the number of iterations of the training process. To solve for $f_t(\cdot)$ from the iteration t to the optimize objective on Eq. (3), second order Taylor expansion is used, thus obtained

$$L^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i (f_t(x_i))^2 \right] + \Omega(f_t) \quad (4)$$

Where $g_i = L'$ is the gradient and $h_i = L''$ is hessian. L' in which the First derivative and L'' is the second derivative of optimize objective.

In Vanilla XGBoost the search for gradients and Hessians is based on the use of binary cross entropy (\mathcal{L}_{bce})[18].

$$\mathcal{L}_{bce} = -\frac{1}{m} \sum_{i=1}^m y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (5)$$

Meanwhile, on the Imbalanced XGBoost proposed by Wang et al. Binary cross entropy is replaced by weighted binary cross entropy (\mathcal{L}_{wbce}) and Focal Loss (\mathcal{L}_F) to solve for imbalance problem [11].

$$\mathcal{L}_{wbce} = -\sum_{i=1}^m \alpha y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (6)$$

$$\mathcal{L}_F = -\sum_{i=1}^m y_i (1 - \hat{y}_i)^\gamma \log(\hat{y}_i) + (1 - y_i) (\hat{y}_i)^\gamma \log(1 - \hat{y}_i) \quad (7)$$

Where α is an imbalance parameter that work for tuning the weight of the class, while γ is a parameter to manage the shape of the curve, the lower the value of γ , the higher the loss and vice-versa. In this case the γ is important aspect to handling the imbalanced class[19].

3.3 Proposed method

Based on weighted binary cross entropy concept, we proposed modified focal loss (\mathcal{L}_{MF}) by adding new imbalance parameter (φ). The φ value is obtained from imbalanced parameter from W-CEL function proposed by whereas[13]

$$\mathcal{L}_{MF} = - \sum_{i=1}^m \varphi y_i(1 - \hat{y}_i)^\gamma \log(\hat{y}) + (1 - y_i)(\hat{y}_i)^\gamma \log(1 - \hat{y}) \quad (8)$$

value φ is obtained with slightly change in the equation from [13]

$$\varphi = \frac{1}{m} \sum_{i=1}^m \beta_i \quad (9)$$

$$\beta_i = \begin{cases} \frac{P + N}{P}, & y = 1 \\ \frac{P + N}{N}, & y = 0 \end{cases} \quad (10)$$

Where P represents data in the positive class and N represents data in the negative class and β_i is parameter of the binary class.

3.4 Evaluation

The performance of the classification methods was evaluated based on accuracy, specificity, recall, and F-1 score. These evaluation metrics were applied because of the relevance in assessing the imbalanced classification problem [4].

The initial evaluation for the classification method is accuracy, but the accuracy does not work well in an imbalance dataset. The evaluation tends to emphasize the dominant class [20]. Precision and recall were the correct binary classification assessment since it provided the consistency of each group. For an imbalanced dataset, MCC was an outstanding assessment since the evaluation consists of True Positive (TP), False Negative (FN), True Negative (TN) and False Positive (FP)[4].

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (11)$$

$$recall = \frac{TP}{TP + FN} \quad (12)$$

$$precision = \frac{TP}{TP + FP} \quad (13)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (14)$$

4. Experiment and result

This experiment was carried out by taking 3 scenarios and the dataset that we used in this experiment came from ULB machine learning group. The better explanation about the dataset is explained in Section 2.

The first scenario used raw data so that data with class "fraud" and class "non-fraud" had a large imbalance. This case was called extreme imbalance. In the second scenario, the raw data was minimized so that the "non-fraud" class will have 10'000 data, but the data for the "fraud" class was not changed. This case was called mild imbalance. The last scenario, we did the same as the second scenario, however the difference is the "non-fraud" class will have 125'000 data.

Fig. 1 provides an overview of the distribution of data using PCA values. It can be seen that the data clearly has gaps, especially on the "0" non-fraud side. There is some data that is included in the outliers, but in this scientific article, we are not focusing on the outliers Based on the methodology proposed in Section 3, the first step was to perform the min-max normalization of the 'amount' feature. After normalizing the features, the next step was to perform feature selection using the coefficient correlation method. Fig. 2 represents the correlation score for each feature. Correlation that has value ($r \geq 0.7$) will be removed. Based on the correlation heat-map in Fig. 2, it can be concluded that the removed features are "V3,V5,V7,V9,V10,V11,V12,V14,V16,V17,V19" Therefore, the initial total of 29 features has changed to 18 features due to 11 reduced features.

Furthermore, the experiment will be divided into two parts, namely the experiment without using hyperparameters for traditional machine learning and using hyperparameters to find values γ . These two trials were divided into 70% training and 30% testing.

Proof of problems in imbalance data can be seen in Table 1. In Table 1, we prove it using several traditional machine learning methods: Logistic Regression, SVM, KNN with $k = \{3,5,7\}$, and Naive Bayes. It can be seen that there is quite a big difference in the value of recall and MCC compared to the accuracy value. This makes the model invalid because the data is biased against the "non-fraud" class. The accuracy in this table is 100% due to the rounding done.

Table 2 is the comparison between the proposed method compared to vanilla XGBoost, Imbalanced XGBoost with weighted binary cross entropy loss, and Imbalanced XGBoost with Focal Loss. In this table, we do not use hyperparameter tuning in the

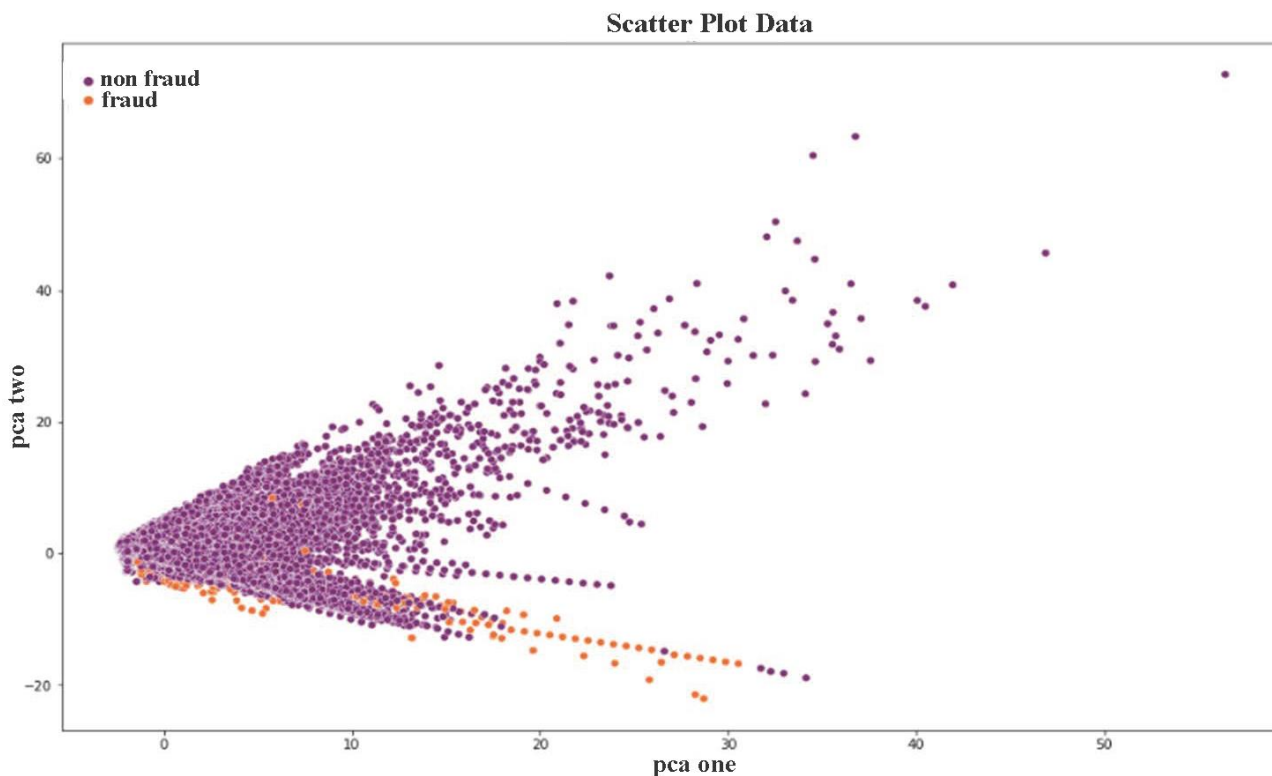


Figure. 1 Distribution of credit card dataset

Table 1. Evaluation of Scenario 1 in traditional machine learning algorithm

	accuracy	Precision	Recall	MCC
Logistic Regression	100	0.76	0.39	0.54
K- Nearest Neighbour (k=3)	100	0.97	0.51	0.7
K- Nearest Neighbour (k=5)	100	0.97	0.47	0.68
K- Nearest Neighbour (k=7)	100	0.94	0.44	0.67
Support Vector Machine	100	0.53	0.53	0.12
Naïve Bayes	97	0.03	0.53	0.12

Table 2. Evaluation of Scenario 1 in XGBoost, imbalanced XGBoost and our proposed

	Accuracy	Precision	Recall	MCC
Vanilla XGBoost [18]	100	0.77	0.39	0.55
Weighted -Imbalanced XGBoost [11]	100	0.92	0.52	0.7
Imbalanced XGBoost with Focal Loss [11]	100	0.97	0.57	0.7
Optimized XGBoost [12]	100	0.94	0.59	0.74
Our Proposed Method	100	0.97	0.56	0.72

XGBoost parameter, instead the parameter values γ and α are tuned using GridSearchCV from scikit-learn.

Based on the results obtained, the proposed method increased the precision results similar to the same precision results as the weighted-Imbalanced XGBoost method with a value of 0.97. Meanwhile, the recall and MCC values of the proposed method were the highest with a recall of 0.59 and 0.72, respectively. Unfortunately the results of the proposed method still did not make this model valid. There was still bias towards the dominant imbalanced data.

To prove that the proposed method was working well, we minimized the amount of data from the non-fraud class. The non-fraud class was reduced from 284'315 data to 10'000 data, we called the data as mild imbalance scenario (scenario 2). We also created another scenario that reduced the non-fraud class to 125'000 data, we called the data as medium imbalance scenario (scenario 3). The data selection was done randomly. The amount of data in the 'fraud' class was not changed. Fig. 3 illustrates the comparison of the data in scenario 1, scenario 2 and scenario 3.

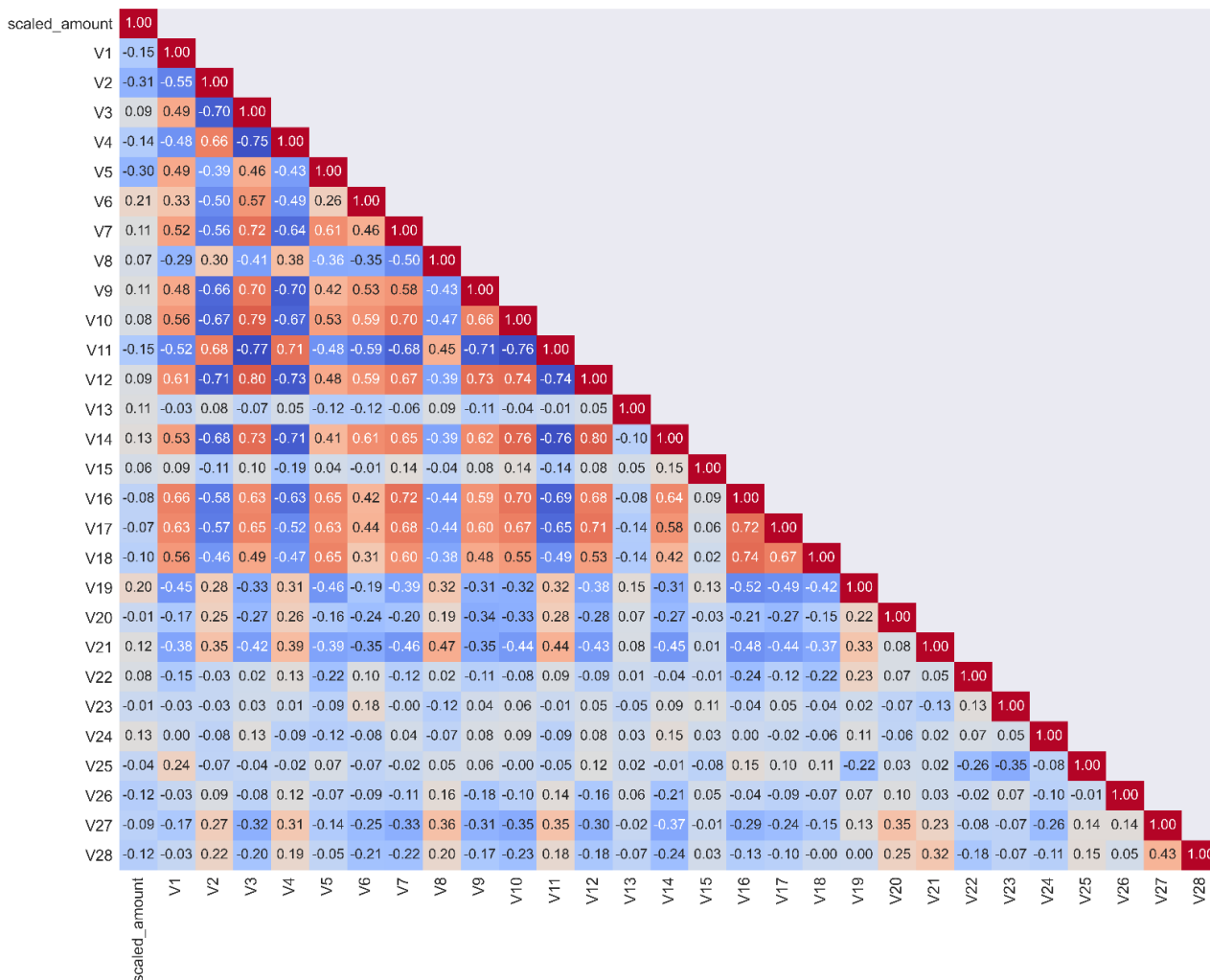


Figure. 2 Correlation score for each feature

Table 3 is the evaluation result using the traditional machine learning method. The results displayed by the model are better but they are still said to be invalid because the values for precision, recall, and MCC differ quite significantly. This statement is attested to in the results of all selected traditional machine learning methods.

Table 4 is the result of the evaluation in scenario 2 of the proposed method which compares vanilla XGBoost with Imbalanced XGBoost. Based on the results obtained, the proposed method succeeded in providing the most valid results because the precision, recall and mcc have close values and the accuracy does not have big gap between other evaluation scores. Meanwhile, the imbalanced XGBoost method with Focal loss can be said to be valid, but there are still discrepancies, especially in the values of precision and recall. Whereas the weighted-Imbalanced XGBoost already gave valid results that were the same as the proposed method, the results of the proposed method were 0.03 points better.

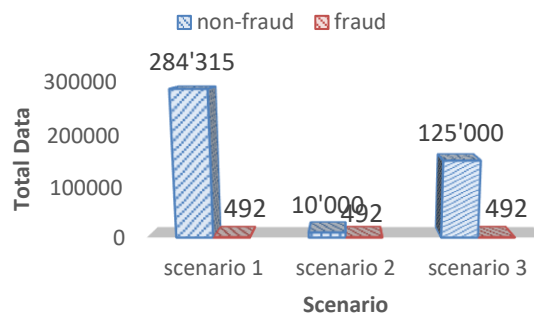


Figure. 3 comparison of the number of data in all scenario

However the Optimized XGBoost, outperform all the other methods, and showed promising result in imbalanced dataset.

Table 5 is the result of the evaluation in scenario 3 of the proposed method which compared between traditional machine learning algorithms, Vanilla XGBoost and Imbalanced XGBoost. In this

Table 3. Evaluation of Scenario 2 in traditional machine learning algorithm

	Accuracy	Precision	Recall	MCC
Logistic Regression	99	0.96	0.76	0.84
K- Nearest Neighbour (k=3)	99	0.98	0.72	0.84
K- Nearest Neighbour (k=5)	98	0.97	0.67	0.8
K- Nearest Neighbour (k=7)	98	0.99	0.62	0.67
Support Vector Machine	99	1	0.8	0.89
Naïve Bayes	95	0.48	0.69	0.55

Table 4. Evaluation of Scenario 2 in XGBoost, imbalanced XGBoost and our proposed

	Accuracy	Precision	Recall	MCC
Vanilla XGBoost [18]	99	0.96	0.77	0.85
Weighted -Imbalanced XGBoost [11]	99	0.93	0.79	0.85
Imbalanced XGBoost with Focal Loss [11]	99	0.84	0.84	0.85
Optimized XGBoost [12]	99	0.92	0.95	0.93
Our Proposed Method	99	0.88	0.87	0.89

Table 5. Evaluation of Scenario 3

	Accuracy	Precision	Recall	MCC
Logistic Regression	100	0.87	0.61	0.73
K- Nearest Neighbour (k=3)	100	0.97	0.71	0.83
K- Nearest Neighbour (k=5)	100	0.97	0.67	0.81
K- Nearest Neighbour (k=7)	100	0.96	0.66	0.8
Support Vector Machine	100	0.81	0.95	0.81
Naïve Bayes	97	0.09	0.65	0.23
Vanilla XGBoost [18]	100	0.93	0.64	0.77
Weighted -Imbalanced XGBoost [11]	100	0.95	0.69	0.81
Imbalanced XGBoost with Focal Loss [11]	100	0.95	0.69	0.81
Optimized XGBoost [12]	100	0.97	0.69	0.82
Our Proposed Method	100	0.97	0.72	0.83

scenario 3, the proposed method showed as one of the best in term of MCC score result. However, similar as the scenario 1, the provided result shows the precision, recall and MCC score differ to one another. It is showing that the proposed method could not create a robust model for imbalanced data in medium-imbalanced dataset.

Based on the experiments given in both scenario 1 and scenario 2, the proposed method can beat traditional machine learning, as well as the imbalanced XGBoost proposed by Wang et al. However, every method that exists does not work to overcome the extreme imbalanced data that is created. However, in mild imbalanced data Imbalanced XGBoost and Modified Imbalanced XGBoost can work well so that it makes the model more valid

5. Conclusion

In this article, we proposed a method for modified Focal loss to solve the imbalanced dataset problem without any pre-processing steps such as sampling and outlier detection. The proposed method took inspiration from Weighted binary cross entropy with

the consideration of imbalanced parameter (ϕ). The formula of weighted score was taken based on W-CEL imbalanced parameter score.

In this experiment, we classified the credit card fraud using ULB machine learning group data. The data obtained from this sampling were 284'807, where transactions with 492 data proved to be fraud, and the rest is a normal transaction. We are also propose 3 different scenarios to proof our proposed method work in several imbalanced environment. The first scenario, we used all the data without any change, the second scenario, we are lowering the normal class into 10'000 data and the last scenario, we are also lowering the normal class into 125'000 data.

The result from the mild-imbalanced data is 0.88, 0.87, 0.89 from the precision, recall, and MCC scores, respectively. This result proves the proposed method works well in creating machine learning model based on XGBoost algorithm that is not bias in mild-imbalanced dataset. Sadly, the proposed method is not work well to create robust machine learning model in medium and extreme imbalanced data. Our assumption for the finding is the weighted on focal

loss are not sufficient method for imbalanced class that too far from each other. This could be a new finding to handling medium and extreme imbalanced data.

Based on this experiment, we conclude that our proposed method can work well, especially on mild imbalanced data, but still needs a lot of attention to the medium and extreme imbalanced data. We are also finding that the optimized XGBoost give an improvement in vanilla XGBoost, and even more robust compared to other methods in mild and extreme dataset. We concluded that finding the optimized parameter for XGBoost would improve the overall performance of the model.

Future research is expected to be able to solve the problems given to extreme and medium imbalanced data, so that it can produce a valid model to use.

Conflicts of Interest

The authors declare no conflict interest

Author Contributions

Conceptualization, methodology, visualization, Formal Analysis and writing original draft have been done by 1st, data preprocessing and visualization have been done by 2nd author, review and editing have been done by 3rd author, and supervision have been done by 4th author.

References

- [1] CNN Indonesia, “Jumlah Kartu Kredit Beredar Naik 2,67 Persen pada Februari,” 15-Apr-2020.
- [2] CNBC, “Best and worst countries at preventing e-commerce fraud”, 21-May-2015.
- [3] Y. Sahin and E. Duman, “Detecting credit card fraud by ANN and logistic regression”, In: *Proc. of 2011 International Symposium on Innovations in Intelligent Systems and Applications*, Istanbul, Turkey, pp. 315–319, 2011.
- [4] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, “Credit card fraud detection using machine learning techniques: A comparative analysis”, In: *Proc. of 2017 International Conference on Computing Networking and Informatics (ICCNi)*, Lagos, pp. 1–9, 2017.
- [5] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, “Credit Card Fraud Detection Using AdaBoost and Majority Voting”, *IEEE Access*, Vol. 6, pp. 14277–14284, 2018.
- [6] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, “Random forest for credit card fraud detection”, In: *Proc. of 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, Zhuhai, pp. 1–6, 2018.
- [7] S. Bagga, A. Goyal, N. Gupta, and A. Goyal, “Credit Card Fraud Detection using Pipeling and Ensemble Learning”, *Procedia Computer Science*, Vol. 173, pp. 104–112, 2020.
- [8] V. N. Dornadula and S. Geetha, “Credit Card Fraud Detection using Machine Learning Algorithms”, *Procedia Computer Science*, Vol. 165, pp. 631–641, 2019.
- [9] A. Muaz, M. Jayabalan, and V. Thiruchelvam, “A Comparison of Data Sampling Techniques for Credit Card Fraud Detection”, *International Journal of Advanced Computer Science and Applications*, Vol. 11, No. 6, pp. 477–485, 2020.
- [10] D. Trisanto, N. Rismawati, M. F. Mulya, and F. I. Kurniadi, “Effectiveness Undersampling Method and Feature Reduction in Credit Card Fraud Detection”, *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 2, pp. 173–181, Apr. 2020.
- [11] C. Wang, C. Deng, and S. Wang, “Imbalance-XGBoost: leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost”, *Pattern Recognition Letters*, Vol. 136, pp. 190–197, 2020.
- [12] C. V. Priscilla and D. P. Prabha, “Influence of Optimizing XGBoost to handle Class Imbalance in Credit Card Fraud Detection”, In: *Proc. of 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, pp. 1309–1315, 2020.
- [13] R. Qin, K. Qiao, L. Wang, L. Zeng, J. Chen, and B. Yan, “Weighted Focal Loss: An Effective Loss Function to Overcome Unbalance Problem of Chest X-ray14”, *IOP Conf. Ser.: Mater. Sci. Eng.*, Vol. 428, Oct. 2018.
- [14] A. D. Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, “Calibrating Probability with Undersampling for Unbalanced Classification”, In: *Proc. of 2015 IEEE Symposium Series on Computational Intelligence*, Cape Town, South Africa, pp. 159–166, 2015.
- [15] H. Tingfei, C. Guangquan, and H. Kuihua, “Using Variational Auto Encoding in Credit Card Fraud Detection”, *IEEE Access*, Vol. 8, pp. 149841–149853, 2020.
- [16] E. Ogasawara, L. C. Martinez, D. de Oliveira, G. Zimbrao, G. L. Pap, and M. Mattoso, “Adaptive Normalization: A novel data normalization approach for non-stationary time series”, In:

Proc. of the 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, Spain, pp. 1–8, 2010.

- [17] E. C. Blessie and E. Karthikeyan, “Sigmis: A Feature Selection Algorithm Using Correlation Based Method”, *Journal of Algorithms & Computational Technology*, Vol. 6, No. 3, pp. 385–394, Sep. 2012.
- [18] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System”, In: *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, Aug. 2016.
- [19] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection”, In: *Proc. of 2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 2999–3007, 2017.
- [20] A. Namvar, M. Siami, F. Rabhi, and M. Naderpour, “Credit risk prediction in an imbalanced social lending environment”, *IJCIS*, Vol. 11, No. 1, p. 925, 2018.