

Fraudulent Transaction Detection in FinTech using Machine Learning Algorithms

Khadija AbdulSattar
College of Information Technology
University of Bahrain
Sakhir, Bahrain
20124209@stu.uob.edu.bh

Mustafa Hammad
College of Information Technology
University of Bahrain
Sakhir, Bahrain
mhammad@uob.edu.bh

Abstract— With the advancement of e-commerce, online transactions purchases using credit and debit cards have drastically increased. This has caused a burst in credit and debit card fraud and has become a profoundly significant global issue. Fraud touches every area of our lives and is a growing concern that effects both businesses and customers. As machine learning techniques provide unique and efficient solutions, they are applicable in various types of problems. Recently, machine learning algorithms have been widely applied as a data mining technique for classification problems. In this paper, a binary classification problem is considered where a transaction can be classified as either fraudulent or legitimate transaction. The goal is to classify the transactions using five different machine learning algorithms. The transaction dataset (Task1 and Task2) is preprocessed, and then SGD, DT, RF, J48 and IBk machine learning classifiers are applied. After applying the classifiers, the results are compared to analyze which classifier performs the best. Based on the experimental results, it is found that the accuracy percentage of all the five classifiers for Task1 and Task2 datasets is ranging between 97.78% to 98.1%, with no major difference. As the dataset is highly imbalanced, the kappa statistic value is also considered. For both datasets, the RF classifier had the greatest value of kappa statistics, whereas SGD and J48 had the lowest value for Task1 and Task2, respectively. Other evaluation metrics were also considered for evaluating the performance of the applied classifiers. Overall, these classifiers achieved similar results for Task2 dataset. As negative Kappa statics and MCC values were obtained, the SGD classifier for Task2 dataset had the worst results in comparison. Based on evaluation criteria such as Kappa statistic and MCC values RF outperformed the others for both the datasets.

Keywords— *classification; credit and debit cards transactions; ML algorithms; imbalanced data; fraud detection*

I. INTRODUCTION

In this digital-era of information, the number of transactions being performed through online purchases using credit and debit cards has drastically increased. Fraudulent transaction detection of debit and credit cards is typically considered as a binary classification problem, where each transaction may either be classified as either legitimate or fraudulent. Thus, fraud detection has become a deeply relevant global issue. Therefore, many techniques have been used by the FinTech industry for fraud detection [1-3].

As machine learning techniques provide unique and efficient solutions, they are applicable in various types of problems. There are a large number of research studies conducted on machine learning as a popular method for classification problems. Machine learning helps in revealing scams in different domains, such as FinTech, Healthcare, and eCommerce. There is great emphasis given in the research studies on data mining and neural network for detecting fraudulent transactions [4][5]. In [6], Bhatia et al. discussed a few detection techniques used in the literature. These techniques include a fusion approach, Bayesian and neural

networks, hidden Markov model, SVM, fuzzy Darwinian detection, kNN, and Naïve Bayes.

Although, fraud detection has a long history, but due to data deficiency and other challenges this area has not been developed much. Due to privacy reasons, the real data of banks customers' transactions needs to be kept confidential, and cannot be revealed. Thus, real transaction data is not easily available for exploration. Moreover, the field names of the dataset are changed for privacy reasons. Many challenges are encountered during fraud detection. Puh and Brkić have discussed some of these challenges in [7]. These challenges include data deficiency, imbalanced data, behavioral variation, cost sensitive problem and evaluation metric.

There are different kinds of frauds found in credit card transactions. In [8], G. and G. Gupta explained a few amongst them. These include identity theft, unauthorized purchases using lost/stolen cards, ATO (Account TakeOver) i.e. unauthorized usage of online accounts, and faking and skimming/counterfeiting cards. They performed the fraud detection in two phases, feature extraction and then classification. According to [9], in 2018 identity theft fraud reports, credit card fraud ranked the first. It is being the most common and popular kind of identity theft. It was also stated that credit card fraud increased by 18.4 percent in 2018 and is still climbing [9]. Therefore, fraud detection and prevention are profoundly a significant global issue.

In this paper, a binary classification problem is considered where the target attribute will identify a transaction as either fraudulent or legitimate. The goal is to classify the transactions using different machine learning classifiers. The transaction dataset will be preprocessed, and then five different machine learning classifiers will be applied. After applying the classifiers, the results will be compared to analyze which classifier performs the best.

The sections covered of this paper are as follows. In section II, the related work is given followed by section III where the research methodology is briefly explained. Section IV describes the machine learning algorithms applied in the research. Then, the details of the dataset used, and its pre-processing techniques are explained in section V. Section VI gives all the details about the evaluation measures selected for the performance evaluation of the applied machine learning algorithms. The experimental results are given in section VII. In section VIII, the paper ends with the conclusion and future work.

II. RELATED WORK

Various techniques have been offered to detect fraudulent transactions using neural networks. In [10], a fraud detection method for credit cards was proposed by Ghosh and Reilly which was based on neural networks. Their proposed system was installed in a bank for fraud detection. Similar to [10], Brause et al. [11] presented a credit card fraud prediction

technique depending on neural networks. Aleskerov et al. [12] proposed a technique called CARDWATCH. This is a system used for similar purpose. It provides a graphical user interface for different commercial databases. This system was based on neural network model and provided very convincing fraud detection rates. Another study [13], proposed a model for fraud detection based on neural networks. This was a confidence-based neural network. In this model, a technology for fraud detection that ensured the accuracy and effectiveness was introduced. A transaction was considered as fraudulent, if it had sufficiently low confidence. A parallel Granular Neural Network (GNN) was developed by Syeda et al. [14] for detecting credit card fraud.

There are many studies about fraud detection prediction using hybrid methods. In [15], Patidar and Sharma used Artificial Neural Networks (ANN) along with the genetic algorithm. Here, particularly back propagation neural network was used in the algorithm. Another study [16], also used a combination of techniques. Shen et al. [16] investigated the efficiency of three techniques and then proposed a framework to choose the best among those for detection. The three techniques, decision tree, neural networks and Logistic Regression (LR), were investigated for detecting fraud. In [17], credit card fraud detection classification models based on ANN and LR were implemented and used. It compared the performance of ANN and LR with real data set. A novel approach for credit card fraud detection was proposed by Panigrahi et al. [18]. This proposed approach is a fusion approach which used Dempster-Shafer theory and Bayesian learning. They compared their approach with other methods. It was found that the performance of their approach evidenced very high positive impact in fraud detection for credit cards. In [19], Şahin and Duman developed a technique based on decision trees and SVM. They applied this technique for detecting credit card fraud to a real data set and compared the performance of the two techniques. In [3], hybrid methods which used AdaBoost and majority voting methods were tested on a credit card data set which was publicly available. Then, the model was evaluated.

Moreover, many studies discussed Hidden Markov Model (HMM) for detecting fraud. One such approach was developed by Srivastava et al. [20]. Their model was trained such that, if the transaction had sufficiently high probability, it was considered as fraudulent. Simultaneously, it was ensured that legitimate transactions were not rejected. They compared their models with other techniques in the literature. Likewise, in [21], Dhok and Bamnote presented HMM for detecting credit card fraud during transactions. The proposed HMM helped in achieving a high accuracy on fraud detection while having low false alarm rate. Khan et al. [22] also used HMM for detecting credit card fraud. The experimental results were presented to prove their model's effectiveness. In [23], an Optimized Multiple Semi-HMM (OMSHMM) was developed to optimize the model parameters. It effectively detected fraud transactions. Furthermore, [24] and [25] also used HMM for detecting credit card fraud.

Some related works also discussed approaches which were not based on machine learning algorithms for fraud detection. One such technique was based on association rules, applied by Sánchez et al. [26]. They proposed its use for preventing fraud and detecting unlawful credit card transactions. Their proposed methodology was applied in some retail companies. In [27], Quah et al. presented a model for detecting credit card

fraud in real-time. Their proposed approach used computational intelligence. It made use of self-organizing map to decipher possible fraud cases. It could also filter and analyze customer behavior and spending patterns. Ganji and Mannem [28] proposed a technique called Stream Outliner Detection algorithm based on Reverse k-Nearest Neighbors (SODRNN). Moreover, Chiu and Tsai [29] offered a method for detecting fraud in credit card transactions based on web services collaborative scheme.

From the previous studies, it is found that a variation of machine learning classifiers has been widely applied for detecting fraud in transactions. Most widely used machine learning techniques were based on ANN, LR, SVM, Hidden Markov Model, kNN and decision trees. The use of machine learning techniques based on hybrid models were also applied. In this research, machine learning algorithms RF, SGD, decision table are also applied along with decision trees (J48), and kNN (IBk) which were already applied in other works. Findings of different ML classifiers are compared to evaluate their performance. A number of evaluation measures are considered in order to get a better insight of different results in comparison.

III. METHODOLOGY

In this experimental research, the transaction dataset called UCSD Data Mining Contest 2009 Dataset is used. This research is emphasized on the classification of transactions into either fraudulent or legitimate. The dataset is preprocessed before applying machine learning classifiers for training and testing. Then the performance of different machine learning algorithms is to be compared on this classification problem using Weka tool. The machine learning algorithms analyzed for the transaction dataset are briefly explained in section IV and then the results of different classifiers are to be analyzed and compared. The performance of each classifier is evaluated by a number of classification evaluation measures. Fig. 1 illustrates the model used for fraudulent transaction detection and research design.

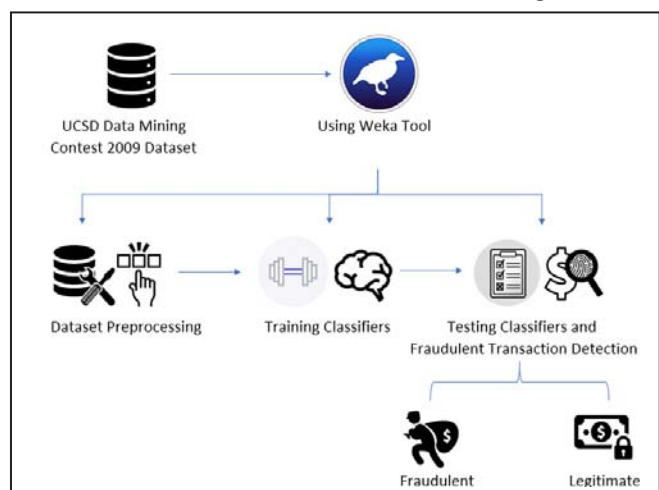


Fig. 1. Model for fraudulent transaction detection and research design

IV. MACHINE LEARNING ALGORITHMS USED

In this study, Weka tool is used as a data mining software for evaluating the datasets. It offers several machine learning algorithms for performing data exploration and data mining tasks. Five machine learning algorithms are selected and the datasets will be tested using these algorithms. These machine learning algorithms are Stochastic Gradient Descent (SGD),

Random Forest (RF), Decision Table (DT), Decision Tree (J48) and Nearest Neighbor algorithm - Instance Based-k (IBk). Each classifier is briefly explained below.

- Stochastic Gradient Descent (SGD): It is an optimization algorithm used for large-scale ML problems, such as natural language processing, as well as, for discriminative learning of linear classifiers [30].
- Random Forest (RF): It is a ML algorithm applied to classification as well as regression problems. Decision trees on data samples are generated. Then, by obtaining the prediction from each generated tree, the best one is chosen based on voting [31].
- Decision Table (DT): It is also a machine learning classification model used for prediction. They are similar to decision trees.
- Decision Tree (J48): It is a supervised learning technique. J48 decision tree classifier is also known as C4.5. It is an algorithm used to generate decision trees. It is used for classification and also known as a statistical classifier [32].
- Nearest Neighbor algorithm - Instance Based-k (IBk): kNN machine learning algorithms is based on supervised learning technique. While classifying an instance using IBk's KNN, its parameter is used for specifying the number of nearest neighbors to use. Then, the outcome is based on majority vote [33].

V. DATA SET AND PRE-PROCESSING

Various kinds of datasets for credit card transaction are available with various transaction attributes. To analyze the classification output of different machine learning algorithms, UCSD Data Mining Contest 2009 Dataset [34] is selected. This dataset is a real dataset of credit card transactions. Fraud detection in credit card transactions is a binary classification problem. Therefore, each credit card transaction belongs to one of the two classes, legitimate or fraudulent. The dataset contains two versions: easy (Task1) and hard (Task2) versions. Both of these versions were used for the evaluation of the classifiers' output.

Table I illustrates fraudulent and legitimate transaction distribution in Task1 and Task2 datasets. It is seen that there is a huge number of fraudulent transactions as compared to legitimate ones. This shows that these datasets are highly imbalanced.

TABLE I. TASK1 AND TASK2 DATASET INSTANCES AND ATTRIBUTES

Dataset details	Dataset version	
	Task1	Task2
Number of instances	94682	100000
Number of attributes (original)	20	20
Number of attributes (after pre-processing)	16	16
Number of fraudulent transactions	2094	2654
Number of legitimate transactions	92588	97346

These datasets were pre-processed by organizing the original dataset, simplifying it and removing all uncommon attributes between the two datasets. As mentioned in Table I, both the original datasets had 20 attributes but after pre-processing 16 attributes were used for analysis. The list of attributes after pre-processing are amount, hour1, state1, field1, field2, hour2, total, flag1, field4, indicator1, indicator2, flag2, flag3, flag4, flag5, and Class. The fields amount, hour1,

hour2, state1, total and class have descriptive column name representing corresponding values for each customer. All the other fields field1, field2, flag1, field4, indicator1, indicator2, flag2, flag3, flag4 and flag5 are anonymized, therefore, are kept as they are. Due to privacy reasons, some attributes are found to be representing the same information as labelled, while all other fields are anonymized. Each version of Task1 and Task2 dataset is further divided into two datasets: the training dataset and the testing dataset. The training dataset has labelled class values, whereas the testing dataset is unlabeled. Henceforth, the training dataset of both the versions Task1 and Task2 are used.

VI. EVALUATION MEASURES

For evaluating the output of the machine learning algorithms, well known evaluation measures are considered. The performance of each classifier is evaluated by classification evaluation measures. The evaluation measures are selected based on their relevance to the problem. Commonly applied evaluation measures in credit card fraud detection problem for assessing the classifier's performance include correctly classified rate, fraud detection rate, Kappa statistic, precision, recall, F-measure, RCC (Receiver Operating Characteristic) area [35] and MCC (Matthews correlation coefficient). The performance of each classifier is evaluated using the below measures:

A. Confusion matrix

Confusion matrix as illustrated in Table II, is generated as a part of the classifier output. It provides TP, TN, FP, and FN values, which is True Positive, True Negative, False Positive and False Negative, respectively. Here, positive class indicates fraudulent and negative class indicates legitimate, hence, the terms TP, TN, FP, and FN are defined as follows:

- TP: the number of fraudulent transactions classified as fraudulent.
- TN: the number of legitimate transactions classified as legitimate.
- FP: the number of legitimate transactions classified as fraudulent.
- FN: the number of fraudulent transactions classified as legitimate.

TABLE II. 2 X 2 CONFUSION MATRIX

Predicted	Actual	
	Fraudulent	Legitimate
Fraudulent	TP	FP
Legitimate	FN	TN

B. Accuracy

Accuracy [36] is the ratio of the number of correctly classified instances to the total number of classified instances. If the value of the ratio is 1, it indicates best accuracy, whereas the value 0 indicates the worst. It is computed as in (1).

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (1)$$

C. Precision

Precision [36] is the ratio of the number of correctly classified positive instances to the total number of positive classified instances. If the value of the ratio is 1, it indicates best precision, whereas, the value 0 indicates the worst. It is computed as in (2).

$$Precision = Confidence = \frac{TP}{TP + FP} \quad (2)$$

D. Recall

Recall [28] is the ratio of the number of classified positive instances to the total number of positive instances. If the value of the ratio is 1, it indicates best recall, whereas, the value 0 indicates the worst. It is computed as in (3).

$$Recall = Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

E. F-measure

F-measure [36] is the weighted harmonic mean of precision and recall of the test, calculated based on the values of recall and precision measures combined in one. It is computed as in (4).

$$F = 2 \left(\frac{Precision \times Recall}{Precision + Recall} \right) \quad (4)$$

F. Kappa statistics

Kappa statistics is an evaluation measure which compares an observed accuracy with an expected accuracy. If the value of Kappa statistic is greater than 0 it indicates better classification. The closer the value to 1 it interprets a better agreement beyond chance. It is computed as in (5).

$$Kappa = \frac{observed\ accuracy - expected\ accuracy}{1 - expected\ accuracy} \quad (5)$$

G. MAE and RMSE

MAE and RMSE are also used for evaluating the performance of a machine learning prediction model. They are used to measure the difference between the values predicted by the model and the actual observed values. If the actual value is a and the predicted value is p then MAE and RMSE are calculated as in (6) and (7).

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - a_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - a_i)^2} \quad (7)$$

H. MCC

Matthews Correlation Coefficient (MCC) [37] is used for the classifier's performance evaluation measures for binary classification problem. It uses TP, FP, TN and FN values. This is one the measures considered for an imbalanced dataset, such as credit card fraud detection dataset. It is a correlation coefficient between the observed and predicted two-class classifications. Its value ranges between -1 to +1, where +1 indicates best prediction, and -1 indicates otherwise. The value 0 indicates no better prediction than chance. It is defined as in (8).

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (8)$$

VII. EXPERIMENTAL RESULTS

When testing the dataset using the above-mentioned classifiers, cross-validation (folds=10) as test option is used. In this test mode, the data in datasets are automatically split between training data and testing data. In each cross-validation fold, 90% of the data is taken for training the model

while the remaining 10% of the data is used for testing. Thus, this process is repeated in each fold.

Fig.2 and Fig. 3 show the percentage of correctly classified instances for both datasets Task1 and Task2 ranges between 97.56% to 98.1%, indicating most of the instances are classified correctly. The number of correctly classified instances indicates the sum of TP and TN values. The number of incorrectly classified instances indicates the sum of FP and FN values. These TP, TN, FP and FN values represent confusion matrices which are given in Table III.

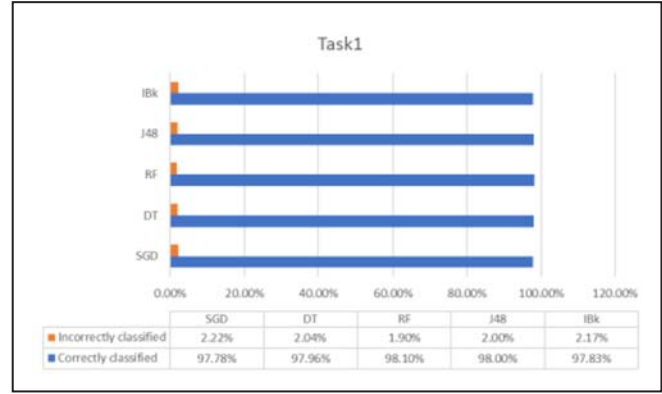


Fig. 2. Comparison between correctly and incorrectly classified transactions of each classifier on Task1 dataset

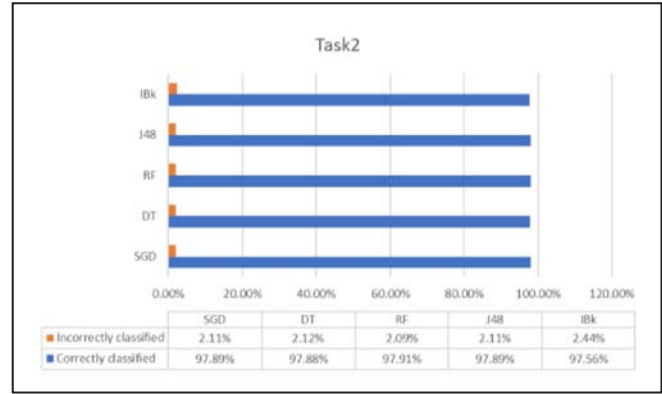


Fig. 3. Comparison between correctly and incorrectly classified transactions of each classifier on Task2 dataset

As mentioned earlier, these datasets are highly imbalanced. Therefore, some other evaluation measures are also needed to be taken under consideration. The accuracy percentage alone cannot be used to conclude the best classifier. One important measure in such datasets is Kappa statistic, if the value is greater than 0 it indicates better classification. The closer the value to 1 interprets a better agreement beyond chance. For Task1 the Kappa statistic values ranges from -0.0001 (SGD classifier) to 0.3381 (RF), indicating SGD performed worst and RF performed the best. Unlike the Task1 Kappa statistic results, Task2 all positive values indicating better performance, the values ranging from 0.3889 (DT) to 0.4657 (RF). Hence, RF outperforms other classifiers in both Task1 and Task2 datasets. Fig. 4 and Fig. 5 show the comparison between Kappa statistic, MAE and RMSE of each classifier SGD, DT, RF, J48 and IBk on Task1 and Task2 datasets, respectively.

The confusion matrix generated as a part of each classifier output is shown in Table III. It further explains the generated results, providing TP, TN, FP and FN values for Task1 and Task2 datasets.

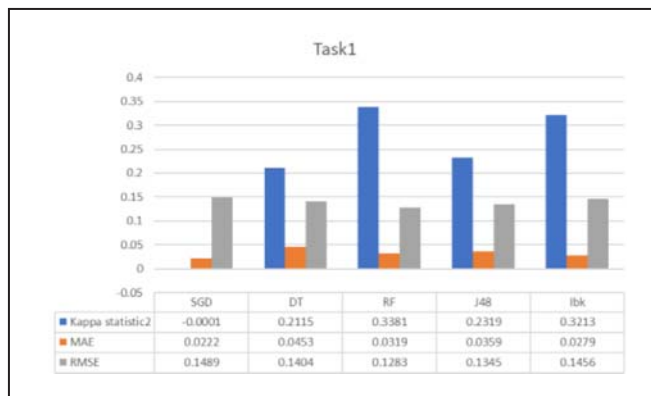


Fig. 4. Comparison between Kappa statistic, MAE and RMSE of each classifier on Task1 dataset

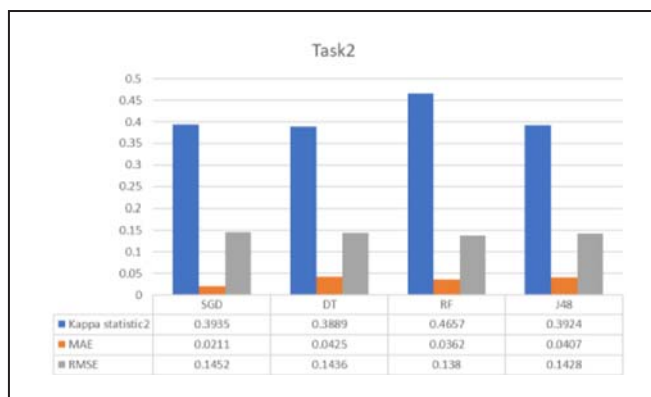


Fig. 5. Comparison between Kappa statistic, MAE and RMSE of each classifier on Task2 dataset

TABLE III. CONFUSION MATRIX OF 5 CLASSIFIERS' OUTPUT FOR TASK1 AND TASK2

Classifier	Predicted	Actual			
		Task1		Task2	
		Fraudulent	Legitimate	Fraudulent	Legitimate
SGD	Fraudulent	92583	5	97185	161
	Legitimate	2094	0	1947	707
DT	Fraudulent	92485	103	97177	169
	Legitimate	1827	267	1955	699
RF	Fraudulent	92413	175	96970	376
	Legitimate	1621	473	1710	944
J48	Fraudulent	92497	91	97187	159
	Legitimate	1800	294	1950	704
IBk	Fraudulent	92124	464	96617	729
	Legitimate	1587	507	1710	944

Table IV show the detailed accuracy by class of all the 5 classifiers used for Task1 and Task2 datasets. Precision, Recall, F-Measure and MCC are well-known evaluation measures. For each classifier, the values the row indicates the weighted average value of the evaluation measures. The weighted average values of precision, recall and F-measure of all the classifiers for both Task1 and Task2 gives strong results with the values ranging from 0.956 to 0.981. The MCC values of Task1 varies for each classifier, with the lowest being -0.001 (SGD classifier) and highest being 0.4 (RF). Based on the MCC values SGD performed worst and RF performed best. Unlike Task1 values of MCC, the Task2 values of MCC are similar for all classifiers ranging between 0.436 to 0.495. Hence, all the classifiers performed almost the same and are

positive values. To be precise the highest value was achieved using RF classifier i.e. 0.495.

TABLE IV. WEIGHTED AVERAGE VALUES OF PRECISION, RECALL F-MEASURE AND MCC OF 5 CLASSIFIERS' OUTPUT FOR TASK1 AND TASK2

Weighted Average for Classifier	Task1				Task2			
	Precision	Recall	F-Measure	MCC	Precision	Recall	F-Measure	MCC
SGD	0.956	0.978	0.967	-0.001	0.976	0.979	0.974	0.459
DT	0.975	0.980	0.973	0.298	0.976	0.979	0.973	0.453
RF	0.977	0.981	0.976	0.400	0.976	0.979	0.976	0.495
J48	0.976	0.980	0.973	0.322	0.976	0.979	0.974	0.458
IBk	0.973	0.978	0.974	0.346	0.972	0.976	0.973	0.436

VIII. CONCLUSION AND FUTURE WORK

With the popularity and upsurge of online transactions, credit card fraud detection has become an essential need and requirement. In this paper a binary classification problem was considered where the target attribute identified a transaction as either fraudulent or legitimate. The goal was to classify the transactions using different machine learning algorithms. The machine learning algorithms applied were SGD, RF, DT, J48 and IBk. For conducting this research, the data mining tool called Weka was used and the datasets were evaluated. For analyzing the classification output of different machine learning algorithms, UCSD Data Mining Contest 2009 Dataset was selected. The transaction dataset was pre-processed, and then the selected machine learning classifiers were applied. After applying the classifiers, the results were compared to analyze which classifier performs the best. Accuracy, precision, recall, F-measure, Kappa statistics, MAE, RMSE, MCC and confusion matrix were used. As the datasets that were used are highly imbalanced, therefore, some other evaluation measures were also needed to be taken under consideration. The accuracy percentage alone cannot be used to conclude the best classifier. Based on the overall experimental results, it is found that for Task1 dataset SGD performed worst and RF performed the best. For Task2 dataset, all the applied classifiers had similar results as per the accuracy. The weighted average values of precision, recall and F-measure of all the classifiers for both Task1 and Task2 gave strong results with the values ranging from 0.956 to 0.981. Based on evaluation criteria such as Kappa statistic and MCC values RF outperformed the others.

This research was limited to classify fraudulent transaction from real dataset that has already happened. Along with fraud detection, fraud prevention is of vital importance. To avoid the damage to happen, it is crucial to prevent such fraudulent transactions in a timely fashion before causing any damage. In the future, this study may further be extended to explore more credit card fraud detections using balanced dataset and by applying feature selection techniques.

REFERENCES

- [1] T. Hastie, R. Tibshirani and J. Friedman, *The elements of statistical learning*. New York, NY: Springer, 2009.
- [2] D. Jensen, "Prospective assessment of ai technologies for fraud detection: A case study", in *AAAI Workshop on AI Approaches to Fraud Detection and Risk Management*, 1997, pp. 34-38.
- [3] K. Randhawa, C. Loo, M. Seera, C. Lim and A. Nandi, "Credit Card Fraud Detection Using AdaBoost and Majority Voting", *IEEE Access*,

- vol. 6, pp. 14277-14284, 2018. Available: 10.1109/access.2018.2806420 [Accessed 22 February 2020].
- [4] S. Benson Edwin Raj and A. Annie Portia, "Analysis on credit card fraud detection methods", *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, 2011. Available: 10.1109/iccet.2011.5762457 [Accessed 22 February 2020].
- [5] S. Maes, K. Tuyls, B. Vanschoenwinkel and B. Manderick, "Credit Card Fraud Detection Using Bayesian and Neural Networks", *First International NAISO Congress on Neuro Fuzzy Technologies, Havana, Cuba*, 2002. [Accessed 22 February 2020].
- [6] S. Bhatia, R. Bajaj and S. Hazari, "Analysis of Credit Card Fraud Detection Techniques", *International Journal of Science and Research (IJSR)*, vol. 5, no. 3, pp. 1302-1307, 2020. [Accessed 22 February 2020].
- [7] M. Puh and L. Brkic, "Detecting Credit Card Fraud Using Selected Machine Learning Algorithms", *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2019. Available: 10.23919/mipro.2019.8757212 [Accessed 22 February 2020].
- [8] G. Gupta, "Analysis of Various Credit Card Fraud Detection Techniques", *International Journal of Computer Science and Mobile Computing*, vol. 8, no. 5, pp. 8-14, 2019. [Accessed 22 February 2020].
- [9] "Credit Card Fraud Statistics", *Shift Credit Card Processing*, 2020. [Online]. Available: <https://shiftprocessing.com/credit-card-fraud-statistics/>. [Accessed: 19- Oct- 2020].
- [10] Ghosh and Reilly, "Credit card fraud detection with a neural-network", *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences HICSS-94*, 1994. Available: 10.1109/hicss.1994.323314 [Accessed 29 February 2020].
- [11] R. Brause, T. Langsdorf and M. Hepp, "Neural data mining for credit card fraud detection", *Proceedings 11th International Conference on Tools with Artificial Intelligence*. Available: 10.1109/tai.1999.809773 [Accessed 29 February 2020].
- [12] E. Aleskerov, B. Freisleben and B. Rao, "CARDWATCH: a neural network based database mining system for credit card fraud detection", *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFER)*. Available: 10.1109/cifer.1997.618940 [Accessed 29 February 2020].
- [13] Tao Guo and Gui-Yang Li, "Neural data mining for credit card fraud detection", *2008 International Conference on Machine Learning and Cybernetics*, 2008. Available: 10.1109/icmlc.2008.4621035 [Accessed 29 February 2020].
- [14] M. Syeda, Yan-Qing Zhang and Yi Pan, "Parallel granular neural networks for fast credit card fraud detection", *2002 IEEE World Congress on Computational Intelligence*. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291). Available: 10.1109/fuzz.2002.1005055 [Accessed 29 February 2020].
- [15] P. R and S. L., "Credit card fraud detection using neural network", *International Journal of Soft Computing and Engineering*, vol. 1, pp. 32-38, 2011. [Accessed 29 February 2020].
- [16] A. Shen, R. Tong and Y. Deng, "Application of Classification Models on Credit Card Fraud Detection", *2007 International Conference on Service Systems and Service Management*, 2007. Available: 10.1109/icsssm.2007.4280163 [Accessed 29 February 2020].
- [17] Y. Sahin and E. Duman, "Detecting credit card fraud by ANN and logistic regression", *2011 International Symposium on Innovations in Intelligent Systems and Applications*, 2011. Available: 10.1109/inista.2011.5946108 [Accessed 29 February 2020].
- [18] S. Panigrahi, A. Kundu, S. Sural and A. Majumdar, "Credit card fraud detection: A fusion approach using Dempster-Shafer theory and Bayesian learning", *Information Fusion*, vol. 10, no. 4, pp. 354-363, 2009. Available: 10.1016/j.inffus.2008.04.001 [Accessed 29 February 2020].
- [19] Y. Sahin and E. Duman, "Detecting credit card fraud by decision trees and support vector machines", in *the International MultiConference of Engineers and Computer Scientists (IMECS)*, Hong Kong, 2011.
- [20] A. Srivastava, A. Kundu, S. Sural and A. Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model", *IEEE Transactions on Dependable and Secure Computing*, vol. 5, no. 1, pp. 37-48, 2008. Available: 10.1109/tdsc.2007.70228 [Accessed 29 February 2020].
- [21] D. Shailesh S. and G. R. Bamnote, "Credit Card Fraud Detection Using Hidden Markov Model", *International Journal of Advanced Research in Computer Science*, vol. 3, no. 3, pp. 816-820, 2012. [Accessed 29 February 2020].
- [22] A. Khan, T. Singh and A. Sinhal, "Implement credit card fraudulent detection system using observation probabilistic in hidden Markov model", *2012 Nirma University International Conference on Engineering (NUiCONE)*, 2012. Available: 10.1109/nuicone.2012.6493206 [Accessed 29 February 2020].
- [23] A. Prakash and C. Chandrasekar, "An Optimized Multiple Semi-Hidden Markov Model for Credit Card Fraud Detection", *Indian Journal of Science and Technology*, vol. 8, no. 2, p. 165, 2015. Available: 10.17485/ijst/2015/v8i2/58081 [Accessed 29 February 2020].
- [24] Bhusari and S. Patil, "Application of Hidden Markov Model in Credit Card Fraud Detection", *International Journal of Distributed and Parallel systems*, vol. 2, no. 6, pp. 203-211, 2011. Available: 10.5121/ijdps.2011.2618 [Accessed 29 February 2020].
- [25] I. Avinash and R. C. Thool, "Credit card fraud detection using Hidden Markov Model and its performance", *International Journal of Advanced Research In Computer Science and Software Engineering (IJARCSSE)*, vol. 3, no. 6, 2020. [Accessed 29 February 2020].
- [26] D. Sánchez, M. Vila, L. Cerda and J. Serrano, "Association rules applied to credit card fraud detection", *Expert Systems with Applications*, vol. 36, no. 2, pp. 3630-3640, 2009. Available: 10.1016/j.eswa.2008.02.001 [Accessed 29 February 2020].
- [27] J. Quah and M. Sriganesh, "Real-time credit card fraud detection using computational intelligence", *Expert Systems with Applications*, vol. 35, no. 4, pp. 1721-1732, 2008. Available: 10.1016/j.eswa.2007.08.093 [Accessed 29 February 2020].
- [28] V. Ratnam Ganji and S. Naga Prasad Mannem, "Credit card fraud detection using anti-k nearest neighbor algorithm", *International Journal on Computer Science and Engineering (IJCSE)*, vol. 4, no. 6, pp. 1035-1039, 2012. [Accessed 29 February 2020].
- [29] Chuang-Cheng Chiu and Chieh-Yuan Tsai, "A Web services-based collaborative scheme for credit card fraud detection", *IEEE International Conference on e-Technology, e-Commerce and e-Service*, 2004. *EEE '04*. 2004. Available: 10.1109/eee.2004.1287306 [Accessed 29 February 2020].
- [30] "1.5. Stochastic Gradient Descent — scikit-learn 0.22.2 documentation", *Scikit-learn.org*, 2020. [Online]. Available: <https://scikit-learn.org/stable/modules/sgd.html>. [Accessed: 04- Apr- 2020].
- [31] "Classification Algorithms - Random Forest - Tutorialspoint", *Tutorialspoint.com*, 2020. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_random_forest.htm. [Accessed: 04- Apr- 2020].
- [32] "What is the algorithm of J48 decision tree for classification ?", *ResearchGate*, 2020. [Online]. Available: https://www.researchgate.net/post/What_is_the_algorithm_of_J48_decision_tree_for_classification. [Accessed: 04- Apr- 2020].
- [33] "Machine Learning - K-Nearest Neighbors (KNN) algorithm - Instance based learning [Gerardnico - The Data Blog]", *Gerardnico.com*, 2020. [Online]. Available: https://gerardnico.com/data_mining/knn. [Accessed: 04- Apr- 2020].
- [34] "Index of /commugrate/data/credit_card", *Cs.purdue.edu*, 2020. [Online]. Available: https://www.cs.purdue.edu/commugrate/data/credit_card/. [Accessed: 04- Apr- 2020].
- [35] M. Nur-E-Arefin and M. Sultan Mahmud, "A Comparative Study of Machine Learning Classifiers for Credit Card Fraud Detection", *IJITIS*, vol. 3, no. 1, pp. 395-406, 2020. Available: 10.1515/IJITIS.2020.3.1.395-406 [Accessed 18 April 2020].
- [36] A. Hammouri, M. Hammad, M. Alnabhan and F. Alsarayrah, "Software Bug Prediction using Machine Learning Approach", *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, 2018. Available: 10.14569/ijacsa.2018.090212 [Accessed 18 April 2020].
- [37] K. Seeja and M. Zareapoor, "FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining", *The Scientific World Journal*, vol. 2014, pp. 1-10, 2014. Available: 10.1155/2014/252797 [Accessed 18 April 2020].