Hindawi
Security and Communication Networks
Volume 2018, Article ID 5483472, 15 pages
https://doi.org/10.1155/2018/5483472

Research Article

An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation

Dahee Choi and Kyungho Lee

Center for Information Security Technologies (CIST), Korea University, Seoul 02841, Republic of Korea

Correspondence should be addressed to Kyungho Lee; kevinlee@korea.ac.kr

Received 7 March 2018; Revised 8 June 2018; Accepted 25 June 2018; Published 25 September 2018

Academic Editor: Ilsun You

Financial fraud under IoT environment refers to the unauthorized use of mobile transaction using mobile platform through identity theft or credit card stealing to obtain money fraudulently. Financial fraud under IoT environment is the fast-growing issue through the emergence of smartphone and online transition services. In the real world, a highly accurate process of financial fraud detection under IoT environment is needed since financial fraud causes financial loss. Therefore, we have surveyed financial fraud methods using machine learning and deep learning methodology, mainly from 2016 to 2018, and proposed a process for accurate fraud detection based on the advantages and limitations of each research. Moreover, our approach proposed the overall process of detecting financial fraud based on machine learning and compared with artificial neural networks approach to detect fraud and process large amounts of financial data. To detect financial fraud and process large amounts of financial data, our proposed process includes feature selection, sampling, and applying supervised and unsupervised algorithms. The final model was validated by the actual financial transaction data occurring in Korea, 2015.

1. Introduction

Financial fraud under IoT environment is the fast-growing issue since the mobile channel can facilitate nearly any type of payments. Due to the rapid increase in mobile commerce and the expansion of the IoT environment, financial fraud in mobile payment has arisen and becomes more common. More than 87 percentage of merchants support either mobile site or a mobile application for online shopping or both [1]. Supporting for mobile wallets also helps to increase the overall use of financial fraud under IoT environment. As a result, mobile payments under IoT platform have reached $194.1 billion in 2017, and mobile proximity payments also increased to $30.2 billion in 2017, compared to $18.7 billion in 2016 [2]. Financial fraud can occur in several ways, but the most frequent case is an unauthorized use of mobile payment via credit card number or certification number. Financial fraud via credit card can be classified into two main categories based on the presence of a credit card: (1) the physical card and (2) the virtual card. To commit credit card fraud with a physical card offline, an attacker has to steal the credit card to carry out the fraudulent transactions. The online credit card fraud that does not require the presence of a credit card mainly occurs under IoT environment, since the payment under IoT environment does not require the presence of a physical payment tool; instead, it needs some information such as card number, expiration date, card verification code, and pin number to make the fraudulent payment. For this reason, financial fraud, which usually takes place under the IoT environment, is the most frequent type of financial fraud that involves taking or modifying credit card information. To address the problem of rapidly arising fraud under IoT environment, financial institutions employ various fraud prevention tools like real-time credit authorization, address verification systems (AVS), card verification value, positive and negative list, etc. [3].

However, existing detection systems depend on defined criteria or learned records, which makes it difficult to detect new attack patterns. Therefore, various methods using machine learning and artificial neural networks have been attempted to capture new financial fraud.Our contributions are as follows:

(a) Research on the various research papers based on the machine learning and artificial neural network techniques

and review of latest detection techniques mainly from 2016 to 2018

(b) Analysis of advantages and limitations for the latest research paper

(c) Model building based on the implementation of reviewed paper and full process experiment based on actual financial transaction dataset

(d) Deriving the result in specific way through validation on each step in the process

(e) Comparison with traditional machine learning and deep learning based on artificial neural networks for fraud detection.

2. Literature Review

We reviewed the latest techniques to detect anomaly and trust relaying in IoT environment. Also, we focused on reviewing the methods and algorithms to detect financial fraud from traditional methods to the latest one. V. Sharma et al. proposed a novel solution in the form of fission computing. The proposed approach relies on the edge-crowd integration for maintenance of trust and preservation of privacy rules in social IoT environment. They performed analysis through numerical simulations by using a safe network and presented a case study on the detection of fake news sources in social IoT environment [4]. Also, a pervasive trust management framework is presented for Pervasive Online Social Networks (POSNs), which is capable of generating high trust value between the users with a lower cost of monitoring [5]. As a solution to identify anomalies in IoT environment, a model was proposed on the basis of cognitive tokens, which provide an Intelligent Sensing Model for Anomalies (ISMA) detection by deliberately inducing faulty data to attract the anomalous users [6]. Van Wyk Hartman suggested automatic network topology detection and fraud detection. If fraud is detected in the distribution network, the system schedules the follow-up and field investigation to investigate and fix the fraud [7]. Also, systems and methods for online fraud detection have been proposed. The front end device generates a first dynamic device identification based on dynamic device characteristics and the back end device generates a second dynamic device identification based on the dynamic device characteristics of the front end device for an authentication [8].

Various learning methods and algorithms have been applied for data analysis and anomaly detection. The learning method of supervised, unsupervised, and artificial neural networks approach has been attempted and a web service-based collaborative scheme for financial fraud detection has been proposed [9, 10]. Also, an efficient fraud detection system which is adaptive to the behavior changes by combining classification and clustering techniques has proposed. The proposed financial fraud detection system is composed of two stages comparing the incoming transaction against the transaction history to identify the anomaly using BOAT algorithm, a scalable algorithm, in the first stage. The false alarm rate suspected anomalies are checked with the fraud history database and decide that the detected anomalies are due to the fraudulent transaction or any short-term change in spending profile in the second stage. BOAT algorithm can also incrementally update a decision tree when the training dataset changes dynamically [11]. The machine learning based research has also been proposed, as a web service-based collaborative scheme for credit card fraud detection [9] . The detection of fraud is based on the genetic algorithm calculation and customer behavior [12], and an efficient financial fraud detection system which is adaptive to the behavior changes by combining classification and clustering techniques, a scalable algorithm named BOAT, has also been proposed [12] . As a traditional method of financial fraud detection, the Dempster–Shafer adder (DSA) based on Dempster–Shafer theory and the use of Bayesian learning research had been proposed. In this research, a transaction conversed with the suspicion score, which can be referred to as the probability of the fraudulent transaction, based on the index in the transaction history database. BLAST and SSAHA algorithm are sequence alignment algorithms and used as the alignment of sequences for an efficient technique to examine the spending behavior of customers [11] . To calculate and predict the probability from the user's existing financial information and to build a multilayer model of program behavior, Hidden Markov Model (HMM) has been proposed. The key idea for applying HMM for anomaly detection is to build a multilayer model of program behaviors using HMMs and various methods [13]. Genetic algorithm calculates and finds critical values which aim to obtain better solutions. During a credit card transaction, the fraud has to be deducted in real time and the number of false alerts is being minimized by using genetic algorithm. The detection of fraud is based on the customer's behavior [14]. Artificial neural network (ANN) is applied for detecting fraud, mainly in the context of supervised classification and it can be used in recognition of characteristics timely and make predictions [12] . CARDWATCH is a database mining system used for credit card fraud detection. The system is based on a neural learning module, interfaced with a variety of commercial databases [15]. The module includes Global Constants Module (GCM), Graphical User Interface Module (GUIM), Database Interface Module (DBIM), Learning Algorithms Library (LAL), and Learning Algorithm Interface Module (LAIM). The proposed system is mentioned as easily extensible and able to work directly on a large variety of commercial databases. Fraud detection with Bayesian Belief Network (BBN) has also proceeded [16]. SODRNN is the reverse K-nearest algorithm for data stream outlier detection. Since SODRNN needs only one pass of scan, it is suitable for the credit card fraud detection of massive data processing [17]. Decision trees and Support Vector Machine (SVM) are a kind of supervised learning method detecting normal transaction and fraud by classifier, which can predict whether the transaction is normal or fraud. Decision tree and SVM are to compare the transaction information with historical profile patterns to predict the probability of being fraudulent for a new transaction [18]. There are also other methods for credit card fraud detection such as fuzzy Darwinian detection which comprises Genetic Programming (GP). Syeda et al. in 2002 proposed fuzzy neural networks which run on parallel machines to speed up the rule production for credit

card fraud detection which was customer-specific. In this technique, the Granular Neural Network (GNN) method that uses fuzzy neural network which is based on knowledge discovery was taken to train the network fast and can be processed for fraud detection in parallel [19]. In APATE approach, intrinsic features derived from the characteristics of incoming transactions such as Recency, Frequency, and Monetary (FRM). The customer spending history and network-based features, by exploiting the network of credit card holders and merchants, are deriving a time-dependent suspiciousness score for each network object [20]. A combined method of decision tree, neural networks, and logistic regression [21], decision trees and Support Vector Machine (SVM) [22], a combined method of decision tree, neural networks (NN), and logistic regression [19], Self-Organizing Map (SOM) combined with Gaussian function [22], and fuzzy logic combined with Self-Organizing Map has been introduced for the financial fraud detection method [23]. A combined method of SVM, random forests, logistic regression [24], Self-Organizing Map Neural Network (SOMNN) [25], genetic algorithm with behavior-based technique, and Hidden Markov Model (HMM) has been attempted [26].

We reviewed the latest financial fraud detection methods using machine learning and deep learning methodology. A total of thirteen recent papers published in 2016 and 2017 were reviewed. This paper mainly focused on the papers with experimental results using existing financial datasets and proving the detection efficiency through the dataset. The method and algorithm applied to the dataset are specified and the validation method is also indicated. We also reviewed the advantages and limitations of each paper. A more detailed review of recent financial detection methods is in Table 1.

3. Model and Methodology

Existing detection systems depend on defined criteria or learned records which make it difficult to detect new attack patterns. To discover new patterns and achieve higher detection accuracy, machine learning methods based on supervised learning and unsupervised learning and deep learning using artificial neural networks have been actively studied. Our proposed research analyzes the most recently used methods in financial fraud studies of machine learning and deep learning from 2016 to 2017. Also, our research implemented both machine learning method and deep learning method to compare the efficiency of detecting financial fraud transactions. In the machine learning process, we applied the unsupervised learning method to discover underlying threats and supervised learning for the accurate classification of fraud transactions under IoT environment. The overall processes of detecting financial fraud include sampling process for class imbalance problem and feature selection process for an accurate model. The previous research papers are mainly related to specific approach such as algorithms, which needs a further step for implementation. Moreover, in applying the methods of machine learning, previous research only used one of the learning methods between supervised and unsupervised learning. However, our research has performed the overall process of financial fraud detection in practical perspective based on supervised and unsupervised learning method. Also, we are proposing a practical method by applying sampling process and feature selection process for solving data unbalanced problem and rapid detection in the real world. Furthermore, our research is expected to be useful for practical use since our experiment has a validation score for each process and is based on real transaction data.

3.1. Machine Learning. Machine learning is a field that machines learn concepts using data, using statistical analysis to predict and classify and input data as an output value. The field of machine learning is divided into supervised learning and unsupervised learning depending on the learning method. Supervised learning predicts the value of input data and is classified with the given label. On the other hand, unsupervised learning is performed in a state where data is not labeled and is often called a clustering process.

The proposed model consists of data preprocessing, sampling, feature selection, application of classification, and clustering algorithm based on machine learning. In this paper, the validation step is performed for each step to verify the effectiveness of proposed financial fraud detection model. In the preprocessing process, data correlation analysis and data cleaning process which cleans the noise data are performed. Also, data transformation, integration, and reduction are included in this process. The following process is the sampling process which evaluates dataset with various ratios for verification through random oversampling and undersampling method. Feature selection process has been performed by the filter-based method. After the feature selection process, clustering process with the proposed algorithm performs and this result is used as a training set in the classification process. By applying supervised algorithms to the previous result, which was derived in the clustering process previously, the higher prediction could be achieved. The model validation process is performed with precision and recall rate through F-measure. The overall structure of the proposed fraud detection system model is as Figure 1.

3.2. Sampling. Imbalanced problem in the data could mislead the detection process to the misclassifying problem and a real transaction dataset of financial transaction usually contains a data imbalanced problem. Previous research has proposed a random minority oversampling method and clustering-based undersampling approach to select the representative data as training data to improve the classification accuracy for minority class [40].

To solve this class imbalanced problem, our research generates various datasets using Synthetic Minority Oversampling Technique (SMOTE) and Random Undersampling (RUS) for the more accurate experiment. SMOTE is an oversampling technique that uses a method of generating arbitrary examples rather than simply oversampling through duplication or replacement [41]. RUS was applied for downsizing the normal transactions by extracting sample data randomly for the class imbalance problem. Since the low ratio of anomalous data might lead to less precise results, our

Table 1: Review of financial fraud detection methods.

Cited Used	Data Description	Applied Method	Validation Method	Advantage	Limitation
[27]	Credit and debit card transactions of the Spanish bank BBVA, from January 2011 to December 2012	Multilayer perceptron (MLP)	True Positive Rate (TPR), Receiver Operating Characteris-tic (ROC curves)	The improvement of accuracy in detecting results by using parenclitic networks reconstruction for feature extraction	Requires the cases where the features are not correlated or not extracted by parenclitic networks
[28]	Transactions of National banking group of Italy, from April 2013 to August 2013	Multi-objective genetic algorithm	TPR, ROC curves	Provide feature selection process via the auto-tuning method	Should be applied to cases other than Banksealer
[29]	UCSD Data Mining Contest 2009 data	Deep neural network (DNN)	Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Errors (MAE), Root Mean Squared Log Error (RMSLE)	Study on the importance of features based on the deep learning method	Does not have accurate experimental explanation process and validation
[30]	Dataset achieved from the second robotic & artificial intelligence festival of Amirkabir University	Decision trees	F-Measure	Ensemble classification is performed using cost-sensitive decision trees in a decision forest framework	Having a class imbalance problem
[31]	German dataset (which has been used in KDD99 competition), Australian credit cards' open dataset	Particle swarm optimization (PSO), Teaching-learning-based optimization (TLBO)	Confusion Matrix (True positive, True negative, False positive, False negative)	Experiment with various datasets	Detection accuracy is relatively low
[32]	Not specific	Linear Regression, Artificial Neural Networks (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Stump, M5P Tree, Decision Table	Normalized Root Mean Squared Error (NRMSE), TPR, F-Measure	A comparative study using various algorithms	Detection accuracy should be increased
[33]	UCI German credit card dataset	SVM	K-fold Cross validation	As Gaussian kernels including RBF are with appropriate regularization, it guarantees a globally optimal predictor which minimizes both the estimation and approximation errors of a classifier	There is no comparison with other algorithms and there is no explanation to verify SVM algorithms superiority to others
[34]	Credit card transaction data from commercial bank in China	Convolutional Neural Networks (CNN), K-Means	F-Measure	Designing a feature called trading entropy based on the latest consumption preferences for each customer and generating synthetic fraudulent samples from real frauds by a cost-based sampling method	Detection accuracy should be increased

Table 1: Continued.

Cited Used	Data Description	Applied Method	Validation Method	Advantage	Limitation
[35]	Banking transaction dataset in Iran	KNN	Accuracy, Re-call, Precision	A novel approach combining K-nearest neighbor, association rules like Apriori algorithm	The validation is not specific and it is difficult to compare the proposed results with other algorithms
[36]	Open dataset: ccFraud	NN, PSO, Auto-associative neural network (AANN), Particle swarm optimization auto-associative neural network (PSOAANN)	MSE, Classification Rate (CR)	Combined parallelization of the auto-associative neural network in the hybrid architecture	Dataset is highly unbalanced and detection accuracy should be increased
[37]	Open dataset: MIT Human Dynamics Lab	SVM, Fuzzy clustering	TPR, FPR, ROC curves	Divide the fraud detection system into two principal modules	Would be better to compare it with more diverse algorithms
[38]	Transactions from a large national bank in Italy, collected from December 2012 to August 2013	Principal component analysis (PCA), DBSCAN	ROC curves	Operate in online processing	Accuracy by validation is not constant
[39]	Not specific	Self- organizing map (SOM)	TPR, FPR	Division of transactions to form an input matrix and ability to be applied to a large complex set	Only have compared to one algorithm

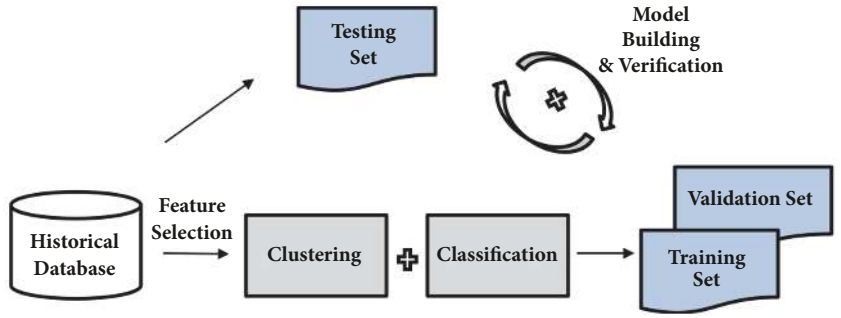


Figure 1: The overall detection process of the proposed process.

research applied both SMOTE and RUS for generating the different ratio of sampling dataset to increase the reliability and accuracy of our proposed research.

3.3. Feature Selection. Feature selection has been proven to be effective and efficient for machine learning problems. The objectives of feature selection include building simpler and more comprehensible models, improving data mining performance such as predictive accuracy and comprehensibility. Also it includes preparing to remove redundancy and irrelevancy for understandable data [42]. Feature selection can be divided into wrapper and filter method. The wrapper method relies on the predictive performance of a predefined learning algorithm to evaluate the selected features. It repeats the searching step and evaluating criteria until desired learning performance is obtained. The drawback of wrapper method is that the search space could be vast and it is relatively more expensive than other methods. Filter method is independent of any learning algorithms and relies on certain characteristics of data to assess the importance of features. Features are scored based on the scores according to the evaluation criteria, and the lowest scored features are removed [43]. For this reason, we applied filter-based feature selection algorithms for feature selection method, which is the fastest and also suitable for practical use. Feature selection based on filter method can be categorized into ranker and the subset selector [44].

In the proposed research, we selected eight subset feature selection algorithms and six ranked feature selection algorithms to select features among existing features. Also, we

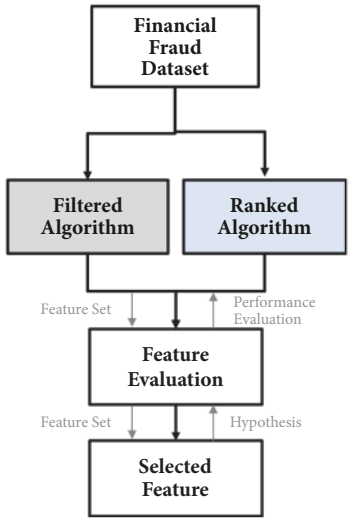


Figure 2: Flowchart of feature selection process.

assigned the score to evaluate features based on the frequency. The ranker algorithms are calculated by weighting the higher ranks. The results of two feature selection algorithms are combined to prioritize the features by selecting features which exceed the parameter in frequency and ranking. Figure 2 is the flowchart of feature selection process proposed in our research.

3.4. Deep Artificial Neural Networks. Deep learning (DL) is a subfield of the machine learning inspired by the structure and function of the brain called artificial neural networks. An artificial intelligence function imitates the working of the human brain in processing data and creating patterns for use in decision-making area, through the capability of unsupervised learning from data that is unstructured or unlabeled. Artificial Neural Networks (ANN) are called neural networks or multilayer perceptrons. A perceptron is a single neuron model that was a precursor to larger neural networks. In neural networks, the predictive capability comes from the hierarchical or multilayered structure of the networks [45]. Also, multilayer perceptron has a neural network with one or more intermediate layers between the input and output layers. Figure 3 is a simple artificial neural network and the middle layer between the input layer and the output layer is called a hidden layer. The network is connected in the direction of the input layer, the hidden layer, and the output layer and is a feedforward network in which there is no direct connection from the output layer to the input layer in each layer. Most multilayer perceptrons can be learned using backpropagation learning algorithms.

3.5. Validation. In machine learning method, which is based on statistics, F-measure is a well-known measurement of model performance between predicted class and actual class using recall and precision. In our research, the F-measure is used to measure the ratio between the actual value and the value that the algorithm detects and predicts [46] and the confusion matrix used to measure the F-measure value is as in Table 2.

4. Experiment

We validated each step to measure the efficiency of the proposed model. Before the feature selection process, the accuracy of each algorithm with raw dataset was measured. After the previous step, the accuracy of each algorithm using the feature extracted through the proposed feature selection method was measured. We used both supervised learning algorithm and unsupervised learning algorithm. In addition to actual datasets, open data were also applied additionally for more accurate verification of our proposed methods.

4.1. Data Description. Our research was conducted based on the actual payment data under IoT environment occurring in Korea, 2015. With the agreement of a major financial institution, provider collected actual financial transaction data for 6 months from June to November. A total of 270,000 pieces of data were extracted from the September data and used as training data. Among data, 21 characteristics are extracted as features (transaction serial number, transaction type, certification date, authentication time, transaction status, telecommunication company, phone number, transaction amount, corporation ID, shop ID, service ID, email hash, IP information, authenticated client version, etc.). For the protection of personal information, key information has been anonymized and data which can identify an individual has been converted to the hash value.

4.2. Modeling Process. In this paper, we aimed to discover hidden patterns by using unsupervised learning and supervised learning for more accurate classification. To design the detection system as to be useful in the operation in the real environment, we proposed the feature selection method that can be applied to the automation system. Therefore, we constructed the system model process by applying the feature selection method on unsupervised learning algorithm firstly and then applied supervised learning algorithm later for accurate classification based on the above experimental results by open dataset and real dataset. The final model validation was performed based on actual financial transaction data in Korea. Also, we compared the final accuracy of the proposed machine learning based detection model and the detection accuracy of models using artificial deep neural networks.

The proposed machine learning based model includes various proportions of the sampling process for application in the real environment and includes an algorithm based automatic feature selection process. In addition, we apply algorithms based on unsupervised learning using selected features and apply algorithms based on supervised learning for more accurate classification. On the other hand, the deep learning model derives the optimum value through the parameter adjustment of the neural networks. Plot (a) in Figure 4 shows the classification accuracy of the UCI German credit card data, by dividing the data before and after the

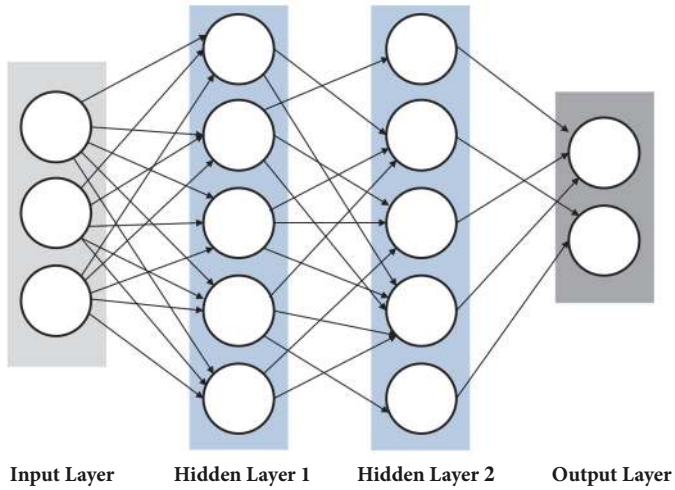


Figure 3: A simple artificial neural network configuration.

Table 2: Confusion matrix.

	Predicted Positive	Predicted Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

feature selection process and applying it to each algorithm to detect abnormal transactions. Applied algorithms are clustering algorithms: EM, simpleK, DensityBased, LVQ, XMeans, FarthestFirst, Hierarchical, and Self-Organizing Map. The purple line indicates the accuracy before the feature selection, and the orange line indicates the accuracy after the feature selection. Plot (b) in Figure 4 shows the F-measure value of classification and also the purple line indicates the value before the feature selection process and the red line indicates the value after the feature selection process.

The results of experiments based on the open dataset show that the F-measure value arises in the majority of algorithms after the proposed feature selection process. The algorithms based on the unsupervised learning have achieved a maximum accuracy improvement of 11.5% and an average of about 11% after the feature selection process in open dataset. Figure 5 is the distribution of accuracy and F-measure value before and after the feature selection process.

Table 3 shows the detailed results of F-measure value in experiments based on the open dataset before and after the feature selection process. Specifically, the proportion of 90:10 ranked the highest in average F-measure value of 0.7475. The proportion of 95:1 ranked second highest in average value of 0.7387 and 99:1 ranked third in the average F-measure value of 0.7131.

The real dataset is difficult to detect due to highly unbalanced data problem. To find the sampling ratio suitable for the financial dataset, various sampling experiments were conducted. Sampling rates were 50:50, 60:40, 70:30, 80:20, 90:10, 95:5, and 99:1. Plot (a) in Figure 6 shows the average of the detection results in accuracy based on clustering algorithms at various sampling ratios. Also, Plot (b) in Figure 6 shows the average of the F-measure value at various sampling ratios. Results indicate that, at dataset, the 95:5 ratio was the most efficient, followed by the 90:10 ratio.

The five algorithms with good detection efficiency among clustering algorithms were selected and experiments were conducted. The selected algorithms were applied to various ratios as described above, and the proposed model was validated using actual financial transaction dataset occurring in Korea, 2015. Details about accuracy and F-measure in various ratios with real dataset occurring in September are as follows in plot (a) and plot (b) in Figure 7. For more accurate validation, we performed validation process with more dataset. Additionally, accuracy and F-measure in various ratios with real dataset occurring in October and November are in Figures 8 and 9.

The accuracy in detection averages of the clustering algorithms in each dataset is in plot (a) in Figure 10, which includes detection values at various sampling ratios. Plot (b) in Figure 10 shows the average of F-measure via clustering algorithms in various sampling ratios.

The final detection is performed through classification algorithms in the process of sampling, feature selection, and clustering. We aimed to discover hidden patterns by using both unsupervised learning and supervised learning. Therefore, we constructed the system model process by applying feature selection and clustering algorithm firstly and then apply classification algorithm later for accurate classification based on the above experimental results. Six types of classification algorithms were used and we divide the results with sampling ratio for more detailed information. For each ratio, the detection rate of the classification algorithm was measured based on the average detection rate of the previous clustering algorithms. Final detection results of classification in the various ratio are as in Figure 11.

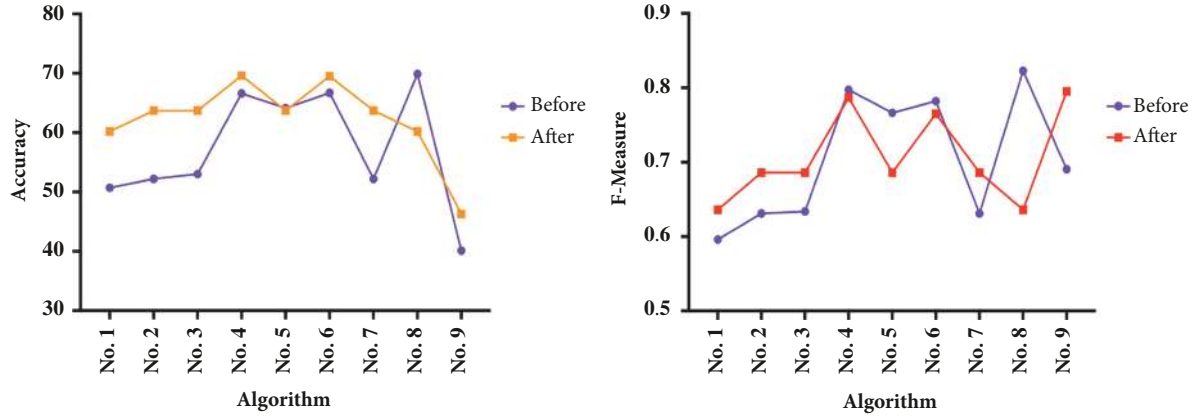


Figure 4: (a) Accuracy before and after the feature selection process. (b) F-measure before and after the feature selection process.

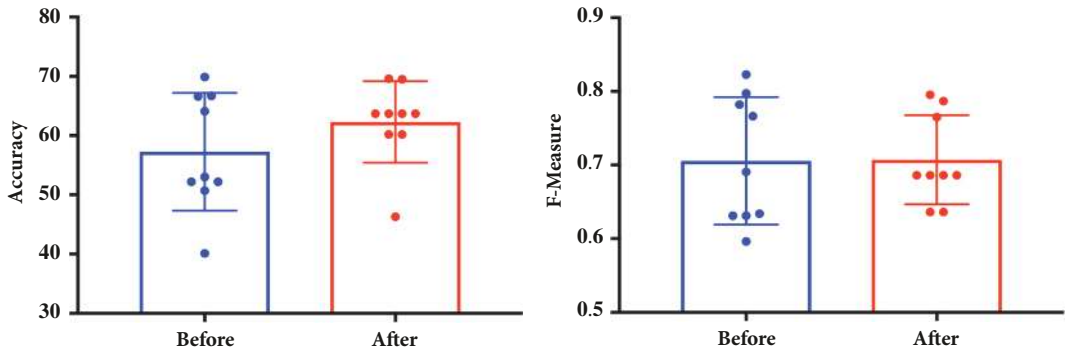


Figure 5: Distribution of accuracy and F-measure value before and after the feature selection process.

Table 4 is the F-measure value of final detection performed through classification algorithms in various sampling ratios.

5. Results

We performed the validation based on the identical actual financial transaction data for machine learning method and artificial neural network. In conclusion, the well-known machine learning method has a higher fraud detection rate than the artificial neural network. By integrating the ratios, the maximum detection rate of the machine learning method was 1, the lowest detection rate was 0.736, and the average detection rate was 0.98618 when all of the algorithms were utilized. The maximum detection rate in all ratios of the artificial neural network was 0.914, the lowest detection rate was 0.651, and the average detection rate was 0.77228. Specific numerical values for each method are shown below.

Results in Figure 12 show the F-measure value of the artificial neural network for detecting financial fraud in various ratios. The ANN achieved an average detection rate of 0.77228 at various ratios; however, it reached a detection rate of 0.914 for each of the 95: 1 and 99: 1 ratios as in Figure 12.

In machine learning model, the experiments were performed from clustering processes such as EM, simple, FarthestFirst, XMeans, and DensityBased algorithms. Classification algorithms such as NaiveBayes, SVM, Regression, OneR, C4.5, and RandomForest were performed and the final result was measured.

In clustering algorithms, EM algorithm reached an average of 0.99862 in fraud detection. DensityBased algorithm ranked second top in fraud detection and reached 0.98788. More details are in Table 5 and Figure 13.

In classification algorithms, Regression reached an average of 0.99971 in fraud detection. Also, RandomForest reached an average of 0.99969 and second top in fraud detection. C4.5 ranked the third tier by reaching 0.99943. More details are in Table 6 and Figure 14.

Comparison with machine learning based classification algorithms and artificial neural networks for accuracy in financial fraud detection is as follows in Figure 15. Six classification algorithms were used for the final classification and compared with the artificial neural network algorithm. We measured the result of the modeling process with F-measure via real dataset. The result classified by well-known machine learning algorithm and the artificial neural network algorithm in various sampling ratios is shown in Figure 15.

Table 3: F-measure of clustering algorithms before and after the feature selection process.

Ratio	Stage	EM	simpleK	FarthestFirst	XMeans	DensityBased
50:50	Before	0.5297	0.4815	0.2692	0.4815	0.4806
	After	0.6134	0.6443	0.6647	0.6443	0.6492
60:40	Before	0.5787	0.5186	0.6869	0.5186	0.5179
	After	0.6319	0.6820	0.7473	0.6820	0.6835
70:30	Before	0.6222	0.6159	0.7527	0.6159	0.6163
	After	0.6453	0.7112	0.8202	0.5400	0.7121
80:20	Before	0.6575	0.6540	0.8085	0.6540	0.6541
	After	0.6540	0.6268	0.8849	0.7178	0.5859
90:10	Before	0.6841	0.6892	0.8590	0.6892	0.6932
	After	0.6610	0.6402	0.9426	0.8126	0.6813
95:1	Before	0.6896	0.7053	0.9028	0.8926	0.7076
	After	0.6732	0.6732	0.9130	0.7608	0.6733
99:1	Before	0.7574	0.7212	0.8328	0.7212	0.7296
	After	0.6829	0.7373	0.7143	0.7538	0.6775
AVG	Before	0.6456	0.6265	0.7307	0.6532	0.6284
	After	0.6516	0.6735	0.8124	0.7016	0.6661

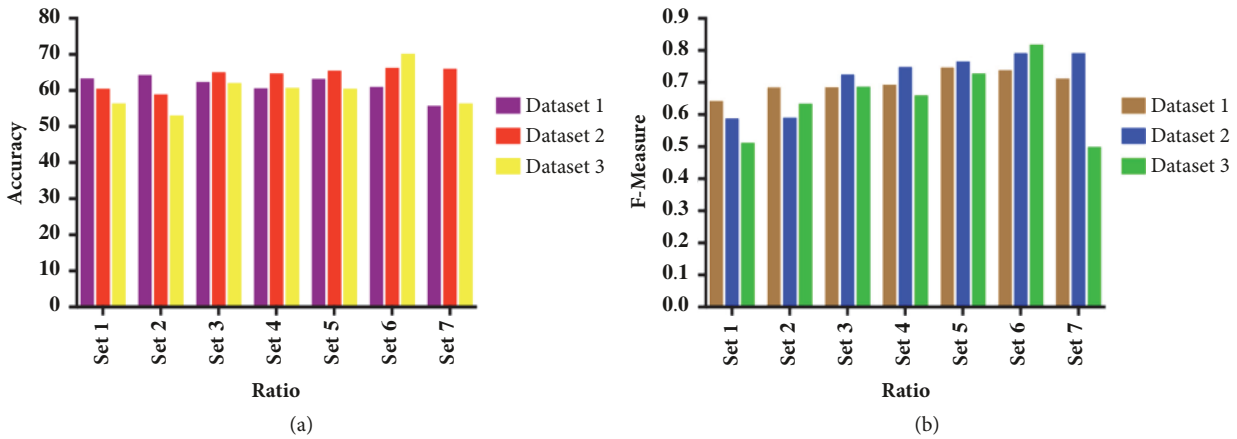


Figure 6: (a) Average detection accuracy in various sampling ratios. (b) Average F-measure in various sampling ratios.

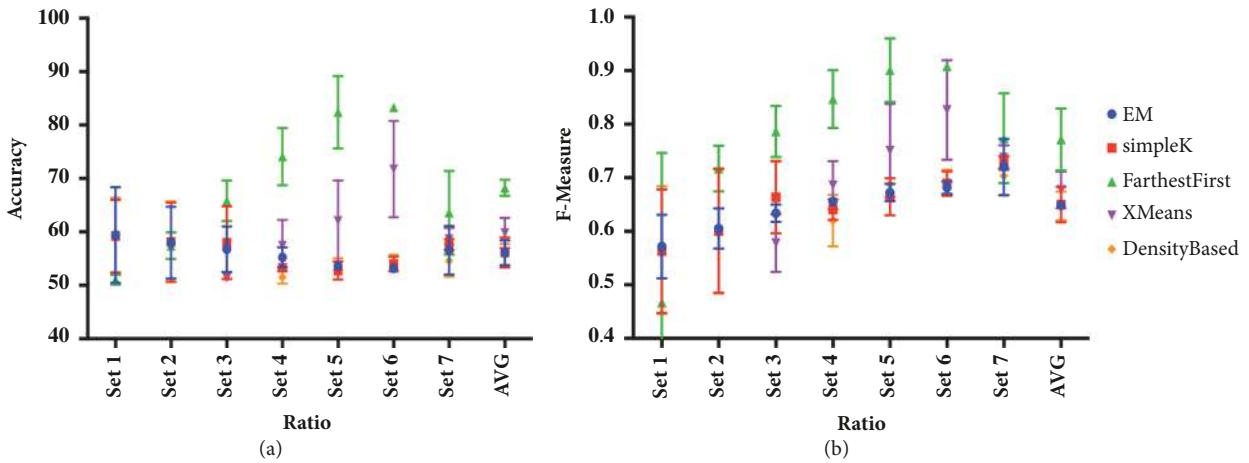


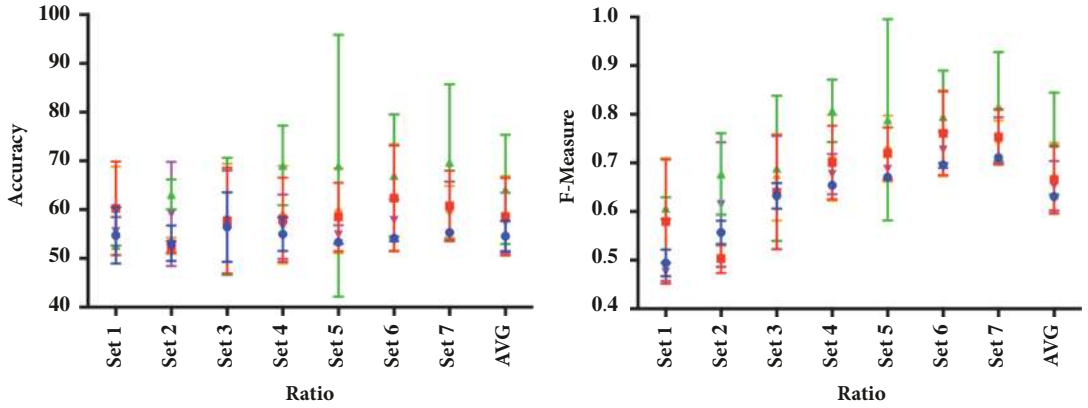
Figure 7: (a) Accuracy in various ratios with real dataset occurring in September. (b) F-measure in various ratios with real dataset occurring in September.

Table 4: F-measure of final classification in various ratios.

	Ratio	Naïve Bayes	SVM	Regression	OneR	C4.5	Random Forest
EM	50:50	1	0.988	0.999	1	1	1
	60:40	1	0.989	0.999	1	1	1
	70:30	0.999	0.999	1	0.983	0.999	1
	80:20	1	0.994	1	1	1	1
	90:10	1	1	1	1	1	1
	95:5	1	1	1	1	1	1
	99:1	0.993	1	1	1	1	1
simpleK	50:50	0.949	0.995	1	0.912	0.997	0.998
	60:40	0.952	0.996	1	0.91	0.998	0.999
	70:30	0.952	0.996	1	0.912	0.999	0.999
	80:20	0.908	0.999	1	0.906	0.999	0.999
	90:10	0.908	1	1	0.898	0.999	1
	95:5	1	1	1	1	1	1
	99:1	0.959	1	1	0.929	1	1
Farthest First	50:50	0.998	0.999	1	0.983	1	1
	60:40	0.998	0.999	1	0.983	0.999	1
	70:30	0.999	1	1	0.983	0.999	1
	80:20	0.999	1	1	0.983	1	1
	90:10	0.999	1	1	0.984	1	1
	95:5	0.995	0.999	1	0.828	1	1
	99:1	0.979	1	1	0.736	1	1
XMeans	50:50	0.949	0.995	1	0.912	0.997	0.998
	60:40	0.952	0.996	1	0.910	0.998	0.999
	70:30	0.996	1	1	0.988	1	1
	80:20	0.949	0.997	1	0.915	1	1
	90:10	0.959	1	1	0.880	1	1
	95:5	0.947	0.999	1	0.913	1	1
	99:1	0.999	1	1	0.908	1	1
Density Based	50:50	0.956	0.989	0.997	0.927	0.998	0.999
	60:40	0.956	0.992	0.998	0.918	0.999	0.999
	70:30	0.955	0.993	0.998	0.919	0.999	0.999
	80:20	0.975	1	1	0.977	1	1
	90:10	0.974	1	0.999	0.975	1	1
	95:5	1	1	1	1	1	1
	99:1	1	1	1	1	1	1

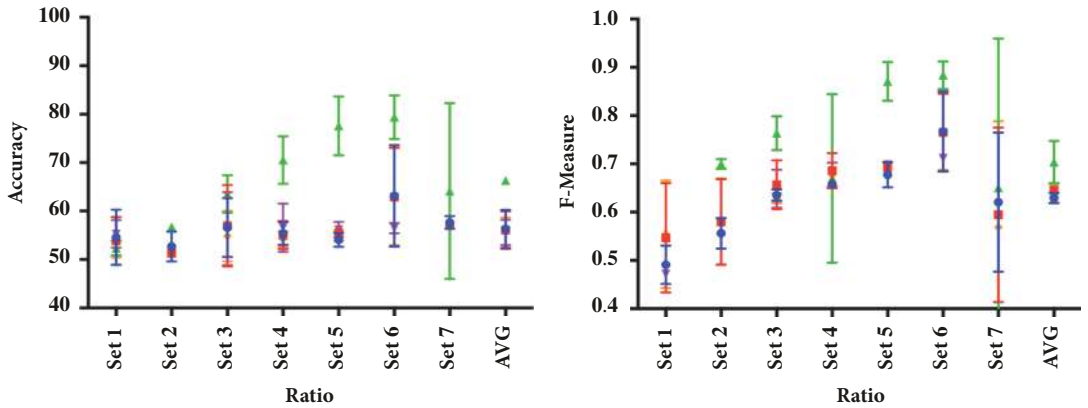
Table 5: Average of clustering algorithms in fraud detection.

	Naïve Bayes	SVM	Regression	OneR	C4.5	Random Forest	AVG
EM	0.99886	0.99571	0.99971	0.99757	0.99986	1	0.99862
simpleK	0.94686	0.99800	1	0.92386	0.99886	0.99929	0.97781
Farthest_F	0.99529	0.99957	1	0.92571	0.99971	1	0.98671
XMeans	0.96443	0.99629	1	0.91800	0.99929	0.99957	0.97990
Density_B	0.97371	0.99754	0.99886	0.95943	0.99943	0.99957	0.98788



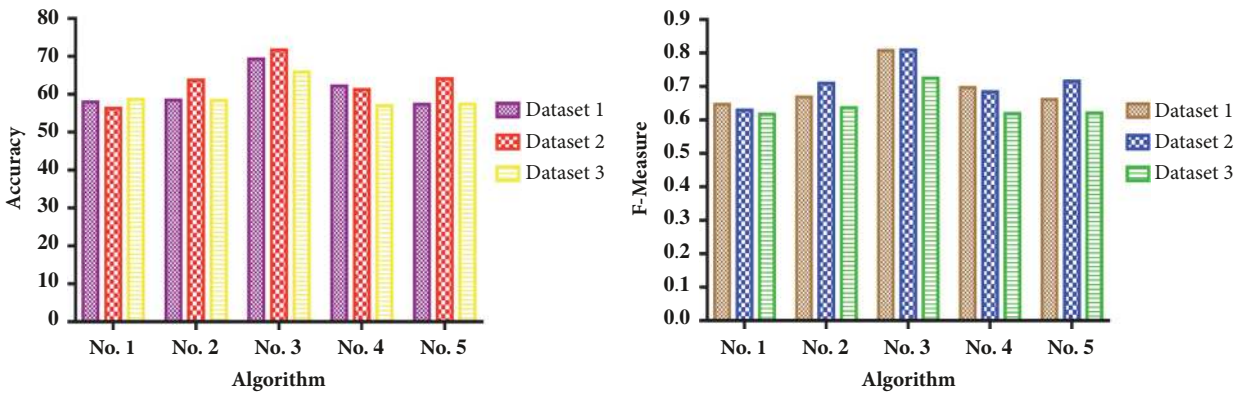
(a) (b)

Figure 8: (a) Accuracy in various ratios with real dataset occurring in October 2015. (b) F-measure in various ratios with real dataset occurring in October 2015.



(a) (b)

Figure 9: (a) Accuracy in various ratios with real dataset occurring in November 2015. (b) F-measure in various ratios with real dataset occurring in November 2015.



(a) (b)

Figure 10: (a) Average of detection accuracy via clustering algorithms in various sampling ratios. (b) Average of F-measure via clustering algorithms in various sampling ratios.

Table 6: Average of classification algorithms in fraud detection.

	EM	simpleK	FarthestFirst	XMeans	DensityBased	AVG
NaiveBayes	0.99886	0.94686	0.99529	0.96443	0.97371	0.97583
SVM	0.99571	0.99800	0.99957	0.99814	0.99629	0.99754
Regression	0.99971	1	1	1	0.99886	0.99971
OneR	0.99757	0.92386	0.92571	0.91800	0.95943	0.94491
C4.5	0.99986	0.99886	0.99971	0.99929	0.99943	0.99943
RandomForest	1	0.99929	1	0.99957	0.99957	0.99969

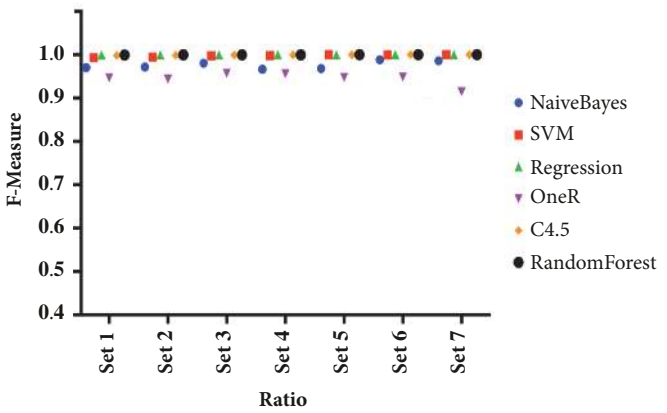


Figure 11: Final detection results of classification algorithms in various ratios.

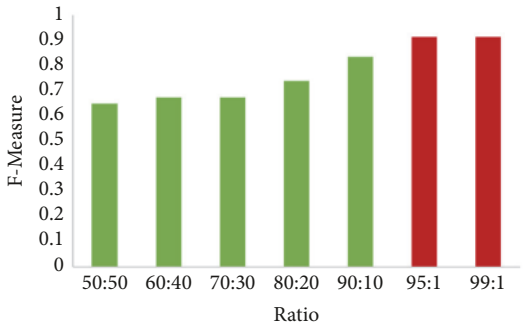


Figure 12: Results of artificial neural network in various ratios for detecting financial fraud.

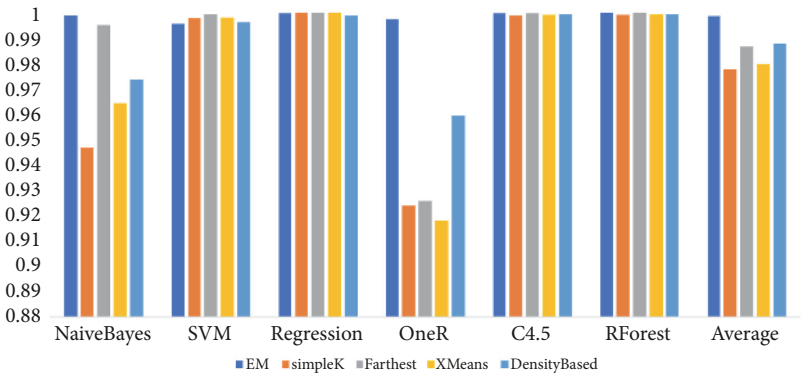


Figure 13: Detection average of clustering algorithms in F-measure.

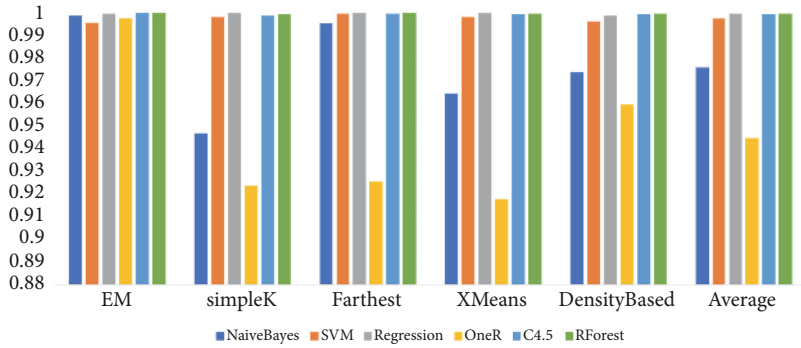


Figure 14: Detection average of classification algorithms in F-measure.

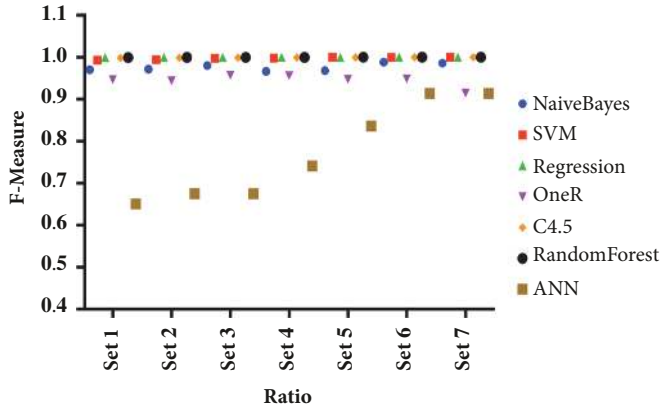


Figure 15: Final financial fraud detection of various algorithms and artificial neural network in various ratios.

6. Conclusions

In this paper, we reviewed the latest financial fraud detection technique using machine learning and artificial neural networks and implemented the experiment based on the real financial data in Korea. The process based on the machine learning method consists of the feature selection process based on the filter method, the clustering process, and the classification process. Experimental results show that machine learning based method has higher detection efficiency than neural networks at various ratios; however, the feature selection process must be performed according to input data. Also, machine learning based process has to verify the optimal combination of clustering algorithms and classification algorithms. Validation of various financial data sets will be performed in the future work. Neural networks reached a particularly high detection accuracy at 95: 1 and 99: 1 ratios, which is nearly similar to the actual ratio in the real world. However, the process takes relatively longer than the machine learning process. In the future work, we aim to improve the accuracy and processing time of the financial fraud process in real time combined with both machine learning based process and deep artificial neural networks.

Data Availability

Our research was conducted based on the actual payment data under IoT environment occurring in Korea, 2016. Since security agreement for the data has been written, it cannot be provided online. The authors are sincerely very sorry for not providing the data online.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] k. corp, "Mobile payments fraud survey report," 2016.

[2] "Javelin strategy and research," 2016.

[3] S. Panigrahi, A. Kundu, S. Sural, and A. K. Majumdar, "Credit card fraud detection: a fusion approach using dempstershafer theory and bayesian learning," *Information Fusion*, vol. 10, no. 4, pp. 354–363, 2009.

[4] V. Sharma et al., "Cooperative trust relaying and privacy preservation via edge-crowdsourcing in social Internet of Things," *Generation Computer Systems*, 2017.

[5] S. Vishal et al., "Computational offloading for efficient trust management in pervasive online social networks using osmotic computing," *IEEE Access*, vol. 5, pp. 5084–5103, 2017.

[6] S. Vishal, Y. Ilsun, and K. Ravinder, "Isma: Intelligent sensing model for anomalies detection in cross platform osns with a case study on iot," *IEEE Access*, vol. 5, pp. 3284–3301, 2017.

[7] V. Wyk and Hartman, "Automatic network topology detection and fraud detection," U.S. Patent No. 9,924,242. 20 Mar. 2018.

[8] A. Favila and P. Shivam, "Systems and methods for online fraud detection," U.S. Patent Application No. 15/236,077, 2018.

[9] C. Phua, V. Lee, K. Smith, and R. Gayler, "A Comprehensive Survey of Data Mining-based Fraud Detection Research," 2010.

[10] C.-C. Chiu and C.-Y. Tsai, "A web services-based collaborative scheme for credit card fraud detection," in *Proceedings of the e-Technology, e-Commerce and e-Service, 2004. EEE'04. 2004 IEEE International Conference on*, IEEE, 2004.

[11] K. K. Sherly and R. Nedunchezhian, "Boat adaptive credit card fraud detection system," in *Proceedings of the Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on*, pp. 1–7, 2010.

[12] K. RamaKalyani and D. UmaDevi, "Fraud detection of credit card payment system by genetic algorithm," *International Journal of Scientific & Engineering Research*, vol. 3, no. 7, 2012.

[13] A. Kundu, S. Panigrahi, S. Sural, and A. K. Majumdar, "Blast-ssaha hybridization for credit card fraud detection," *IEEE Transactions on Dependable and secure Computing*, vol. 6, no. 4, pp. 309–315, 2009.

[14] 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.

[15] Nune, G. Kumar, and P. Vasanth Sena, "Novel Artificial Neural Networks and Logistic Approach for Detecting Credit Card Deceit," *International Journal of Computer Science and Network Security (IJCSNS)*, 2013.

[16] E. Aleskerov, B. Freisleben, and R. Bharat, "Cardwatch: A neural network based database mining system for credit card fraud detection," in *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)*, 1997.

[17] S. Maes et al., "Credit card fraud detection using Bayesian and neural networks," in *Proceedings of the 1st International Naiso Congress on Neuro Fuzzy Technologies*, 2002.

[18] V. R. Ganji and S. N. P. Mannem, "Credit card fraud detection using anti-k nearest neighbor algorithm," *International Journal on Computer Science and Engineering*, vol. 4, no. 6, 2012.

[19] Y. G. Sahin and E. Duman, "Detecting credit card fraud by decision trees and support vector machines," *Proceedings of the International MultiConference of Engineers and Computer Scientists 2011*, 2011.

[20] M. Zareapoor, K. R. Seeja, and M. Afshar Alam, "Analysis on credit card fraud detection techniques: based on certain design criteria," *International Journal of Computer Applications*, vol. 52, no. 3, 2012.

[21] V. V. Vlasselaer et al., "APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions," *Decision Support Systems*, vol. 75, 2015.

[22] A. Shen, R. Tong, and Y. Deng, "Application of classification models on credit card fraud detection," in *Proceeding of the 2007 International Conference on Service Systems and Service Management*, IEEE, 2007.

[23] "2007 International Conference on Service Systems and Service Management," pp. 1–4, 2007.

[24] Y. Zhang, F. You, and H. Liu, "Behavior-based credit card fraud detecting model," in *Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC*, pp. 855–858, 2009.

[25] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, vol. 50, no. 3, pp. 602–613, 2011.

[26] F. N. Ogwueleka, "Data mining application in credit card fraud detection system," *Journal of Engineering Science and Technology*, vol. 6, no. 3, pp. 311–322, 2011.

[27] Z. Massimiliano et al., "Credit card fraud detection through parenclitic network analysis," *Complexity*, 2017.

[28] M. Carminati, L. Valentini, and S. Zanero, "A Supervised Auto-Tuning Approach for a Banking Fraud Detection System," in *Proceeding of the International Conference on Cyber Security Cryptography and Machine Learning*, Springer, Cham, Switzerland, 2017.

[29] P. Yamini, "Credit Card Fraud Detection using Deep Learning," *International Journal of Advanced Research in Computer Science*, 2017.

[30] F. Fadaei Noghani and M. Moattar, "Ensemble Classification and Extended Feature Selection for Credit Card Fraud Detection," *Journal of AI and Data Mining*, vol. 5, no. 2, pp. 235–243, 2017.

[31] G. Maryam and S. Mohammad Abadeh, "Fraud Detection of Credit Cards Using Neuro-fuzzy Approach Based on TLBO and PSO Algorithms," *Journal of Computer & Robotics*, vol. 10, no. 2, pp. 57–68, 2017.

[32] M. Sorkun Cihan and T. Toraman, "Fraud Detection on Financial Statements Using Data Mining Techniques," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 5, no. 3, pp. 132–134, 2017.

[33] M. Kamboj and G. Shankey, "Credit Card Fraud Detection and False Alarms Reduction using Support Vector Machines," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 2, no. 4, 2016.

[34] F. Kang et al., "Credit Card Fraud Detection Using Convolutional Neural Networks," in *Proceeding of the International Conference on Neural Information Processing*, Springer International Publishing, 2016.

[35] M. Khodabakhshi and M. Fartash, "Fraud Detection in Banking Using Knn (K-Nearest Neighbor) Algorithm," *International Conference on Research in Science and Technology*, 2016.

[36] S. Kamaruddin and R. Vadlamani, "Credit Card Fraud Detection using Big Data Analytics: Use of PSOAANN based One-Class Classification," in *Proceeding of the International Conference on Informatics and Analytics*, ACM, 2016.

[37] S. Sharmila and S. Panigrahi, "Use of fuzzy clustering and support vector machine for detecting fraud in mobile telecommunication networks," *International Journal of Security and Networks*, vol. 11, no. 1-2, pp. 3–11, 2016.

[38] M. Carminati et al., "BankSealer: A decision support system for online banking fraud analysis and investigation," *Computers & Security*, vol. 53, pp. 175–186, 2015.

[39] M. Bansal and Suman, "Credit card fraud detection using self organised map," *International Journal of Information & Computation Technology*, vol. 4, no. 13, pp. 1343–1348, 2014.

[40] S.-J. Yen and Y.-S. Lee, "Cluster-based under-sampling approaches for imbalanced data distributions," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5718–5727, 2009.

[41] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-smote: a new over-sampling method in imbalanced data sets learning," *Advances in intelligent computing*, pp. 878–887, 2005.

[42] K. E. P. Baksai, *Feature Selection to Detect Patterns in Supervised and Semi Supervised Scenarios, Ph.D. thesis*, Pontificia Universidad Católica de Chile, 2010.

[43] L. Jundong et al., "Feature selection: A data perspective," *ACM Computing Surveys (CSUR)*, vol. 94, 2017.

[44] F. Bagherzadeh-Khiabani et al., "A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results," *Journal of clinical epidemiology*, vol. 71, pp. 76–85, 2016.

[45] C. François, "Deep Learning with Python," 2017.

[46] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.