Support Vector Machine with Information Gain Based Classification for Credit Card Fraud Detection System

Kannan Poongodi and Dhananjay Kumar
Department of Information Technology, Anna University, MIT Campus, Chennai, India

Abstract: *In the credit card industry, fraud is one of the major issues to handle as sometimes the genuine credit card customers may get misclassified as fraudulent and vice-versa. Several detection systems have been developed but the complexity of these systems along with accuracy and precision limits its usefulness in fraud detection applications. In this paper, a new methodology Support Vector Machine with Information Gain (SVMIG) to improve the accuracy of identifying the fraudulent transactions with high true positive rate for the detection of frauds in credit card is proposed. In SVMIG, the min-max normalization is used to normalize the attributes and the feature set of the attributes are reduced by using information gain based attribute selection. Further, the Apriori algorithm is used to select the frequent attribute set and to reduce the candidate's itemset size while detecting fraud. The experimental results suggest that the proposed algorithm achieves 94.102% higher accuracy on the standard dataset compared to the existing Bayesian and random forest based approaches for a large sample size in dealing with legal and fraudulent transactions.*

Keywords: *Apriori algorithm, credit card fraud detection, information gain, support vector machine.*

Received March 5, 2020; accepted September 7, 2020
https://doi.org/10.34028/iajit/18/2/8

1. Introduction

The credit card fraud is a growing problem and has increased substantially due to the rise of online purchases resulting in fraudulent activities. Banks are facing big challenge to detect credit card fraud in providing secure transactions to its customers. Hence, designing an efficient credit card fraud detection system to minimize losses has become crucial for all banks issuing the cards [16, 25]. The fraud can happen in numerous ways by means of stealing the physical card, creating fake cards or by means of skimming tools. When used at any Point Of Sale (POS) like online site, restaurant, petrol bunk etc., the name, Card Verification Value (CVV) and expiry date are noted by the scammers and also deploy malware to steal the personal data from various sources. The credit card frauds are of different types [8] and can be detected using decision trees, genetic algorithms, artificial neural networks, fuzzy logic etc., [1, 13]. In e-commerce business, fraudulent transactions appearing as legitimate ones is a major problem and hence designing an effective credit card fraud detection mechanism is quite challenging [11].

The Hidden Markov Model (HMM) can be used in detecting the credit card fraud where the model is trained initially based on the normal behavior of a cardholder. If the trained HMM does not accept an incoming transaction, it is considered as fraudulent [23]. The consumer buying behavior could help detect credit card fraud in identifying the fraudulent transactions. The buying pattern of customers prior to each transaction is aggregated in determining the legitimate and fraudulent transactions [14]. The machine learning techniques [21] can be used to detect the credit card fraud. The algorithms like Naïve Bayes, Random Forest, Support Vector Machine (SVM), AdaBoost, Majority Voting, etc. helps in detecting the credit card fraud.

In this paper, fraud detection is performed using SVM with Information Gain (SVMIG) to identify the fraud in the credit card. The information gain based attribute selection is used to normalize the attributes and relevant features are selected. The SVMIG is applied on a credit card dataset followed by the process of discretization to reduce the attribute intervals, min-max normalization to normalize the data, attribute selection to select the better features, frequent itemset mining using Apriori algorithm and a pruning process is performed to reduce the size of the candidates. Finally, SVMIG is carried out to detect the credit card fraud and the classification results determine the legal and fraudulent transactions.

The paper is organized as follows. Studies about various credit card fraud detection techniques are described in section 2. Detection of frauds in credit card using SVM with information gain based classification is presented in section 3 and its results are discussed in section 4. Conclusion and future work is presented in section 5.

2. Literature Survey

Behera and Panigrahi [3] proposed a credit card fraud detection method using hybrid integration of fuzzy clustering and neural network to reduce the misclassification rate. A three phase approach is used to detect the credit card fraud. In the first phase user authentication and card details are verified, while behavioral analysis using fuzzy c-means clustering is carried out in the second phase and finally a suspicion score is calculated to group the transactions into genuine or fraudulent or suspicious. Razooqi *et al*. [22] proposed a system employing fuzzy logic method to detect fraud in credit card by identifying a transaction as legitimate or fraudulent. The data is pre-processed based on the user's behavior of previous transactions.

Wang *et al*. [29] have presented a distributed deep learning model to detect credit card fraud in preserving the privacy of bank's customers that helps in identifying a transaction. The privacy is assured using ternary gradients and applied on a real-world credit card fraud detection dataset. De Sa *et al*. [7] developed Fraud-Bayesian Network Classifier (Fraud-BNC) algorithm to detect credit card fraud automatically by a Hyper-Heuristic Evolutionary Algorithm (HHEA). The Fraud-BNC is applied on a dataset received from a Brazilian online payment service PagSeguro can solve the classification problems.

Van *et al*. [27] developed an approach to automatically detect the credit card fraud in online stores based on the spending behavior of customers. The framework is modeled around intrinsic and network-based features of credit card holder's transactions. The system can handle individual fraud and tested on a company dataset.

The increasing usage of credit cards for electronic payments is vulnerable to credit card fraud. Darwish [5] developed a credit card fraud detection system by a fusion of k-means and Artificial Bee Colony (ABC) algorithm to increase the classification accuracy of legitimate and fraudulent transactions. The designed system filters the dataset based on the customers profile parameters to determine the transaction is genuine or fraudulent. Maes *et al*. [17] designed an automated fraud detection system applying machine learning approach to detect the credit card fraud. An integration of Bayesian and artificial neural network machine learning techniques are used to determine the credit card fraud on financial data.

Jiang *et al*. [15] proposed a fraud detection technique to identify the transaction fraud in online shopping. The methodology identifies the fraud using the aggregation strategy and feedback mechanism. Data pre-processing is carried out by dividing the card holders into different groups. A sliding window algorithm is used to aggregate the transactions followed by feature extraction. The classifier is then trained to determine the incoming transaction as legitimate or fraudulent.

Fu *et al*. [12] proposed a fraud detection system using Convolutional Neural Network (CNN) to trace the behavioral fraud from labeled data. The fraud detection framework consists of training and prediction phases. The training phase operates in offline mode while the prediction phase is online and capable of testing the incoming transaction as legal or fraud. A family of neural networks, self-organizing map [18, 20, 31] can detect credit card fraud based on customer behavior. The self-organizing maps carry out optimal classification of transactions in determining the fraud.

Duman and Ozcelik [10] designed a method integrating the genetic algorithm and scatter search to improve the credit card fraud detection system used in a bank. A score is assigned to each transaction and based on the score the transactions are classified as legitimate or fraudulent. The typical objective of any fraud detection system is to minimize the misclassification cost and it is applied on real data. Teh *et al*. [24] proposed a fraud detection technique based on customer spending behavior which can be used to determine the credit card fraud. The customer profile built on three attributes namely time, amount, and geographical location can identify the transaction as fraudulent based on the spending behavior of the customer. Wang and Han [28] designed a model to predict the credit card fraud based on cluster analysis and integrated support vector machine. The original data is adjusted using k-means clustering algorithm and to improve the system efficiency, integrated learning is implemented using AdaBoost algorithm. Finally, classification is performed using integrated SVM to predict the credit card fraud.

3. Proposed Methodology

The SVMIG handles the process of discretization, min-max normalization, attribute selection, frequent itemset mining and SVM with information gain based classification for credit card fraud detection as depicted in Figure 1. In SVMIG, the discretization process is used to reduce the attributes intervals. As a result of discretization, the min-max normalization process receives the reduced attributes intervals as input. The normalization process decomposes the attributes values into smaller size. The smaller size attributes are selected using the information gain based feature selection algorithm. The low values of information gain are used to determine the credit card frauds. Attributes with high information gain determines the legal. The frequent itemsets are extracted using the Apriori algorithm and pruning is performed to reduce the candidate's itemset size. The frequent itemsets are the input to the SVM with information gain based classification to detect the fraud.

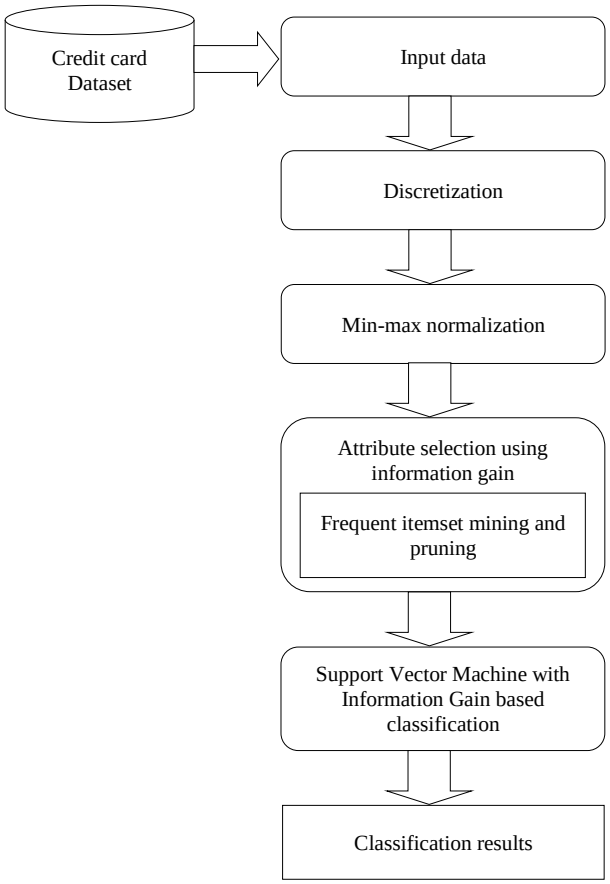

Figure 1. The architecture of SVMIG based classification.

3.1. Discretization

Preprocessing is carried out using discretization and the attribute intervals are reduced by using discretization. The finite number of intervals is transformed into continuous features as a result of discretization. A numerical discrete value is associated with each interval. Data mining modeling performance is enhanced by using discrete values in an appropriate manner and it is very useful in defining frauds in credit card detection system's accuracy. Discretization measure includes the entropy and for a specified attribute *a*, conditional entropy of decision *d* is expressed as $H(d|a)$ and given by,

$$H(d|a) = -\sum_{j=1}^{m}(p_j)\, p(d1|d2)\, logp(d1|d2) \quad (1)$$

Where *p* is the probability, *a* is the attribute, and d_1, d_2 are the decision variables. The two fundamental criteria to calculate the discretization measure based on entropy are information gain associated with attribute *a* and information gain ratio.

3.2. Min–max normalization

The attribute values are decomposed by the process of normalization. Normalization converts the attributes to smaller size and transforms the complex database to a simple one. To test the relation between individuals, sequence of rules is used by normalization. These rules make the normalization to be extended to a greater degree. Min-max normalization is performed after discretization. Before training and testing the data, normalization takes place on the entire data. This is used to ensure the avoidance of data. Linear transformations of original data take place in min-max algorithm. In the default of credit card clients dataset, minimum and maximum value of variables are represented by x_{min} and x_{max}. In min-max algorithm, to map a value v to a value v' is given by,

$$v' = \frac{v - x_{min}}{x_{max} - x_{min}}\,(new\_x_{max} - new\_x_{min}) + new\_x_{min} \quad (2)$$

The min-max algorithm uses linear mapping to scale a variable in training the samples in the interval [x_{min}, x_{max}] to [0, 1] or [-1, 1]. The computation procedure for min-max normalization is given in Algorithm 1.

Algorithm 1: Min-max normalization algorithm

Input: Attributes
Output: Scaled values
1. Begin
2. Set the upper and the lower limit
3. Compute the values of x_{min} *and* x_{max}
4: For each attribute
5. Compute the values of $v' = \frac{v - x_{min}}{x_{max} - x_{min}}\,(new\_x_{max} - new\_x_{min}) + new\_x_{min}$
6. End for
7. End

3.3. Attribute Selection using Information gain

The input data in credit card have several irrelevant attributes and hence selecting the important attributes is time consuming and a difficult process. The most effective feature selection algorithm is information gain based feature selection algorithm and in this case, low value of IG contributes in identifying the credit card frauds.

Dependency between class labels and features are measured using information gain where training data are selected based on information gain. Attributes with high information gain are selected for this purpose and its maximum value corresponds to 1. The information gain is computed separately on every selected feature.

For an attribute *A* and class *C*, entropy of class attribute is given by,

$$H(C) = -\sum_{c \in C} P(c) log\, P(c) \quad (3)$$

The credit card attribute's information entropy with respect to class is defined as,

$$H(C|A) = -\sum_{a \in A} P(a) \sum_{c \in C} P(c|a) log\, P(c| \quad (4)$$

The decrease in entropy provides more class information given by the attribute and it is called as information gain. Depending on information gain of an attribute as well as class, every attribute is assigned with a score A_i.

$$IG_i = H(C) - H(C_i|A) \quad (5)$$

$$IG_i = H(A_i) - H(A_i|C) \quad (6)$$

$$IG_i = H(A_i) + H(C) - H(A_i, C) \quad (7)$$

The information gain has a connection between individual features in the credit card fraud detection. For each category, prediction is measured by information gain using attributes absence and presence. The information gain is given by,

$$IG(t) = -\sum_{i=1}^{C} P(C_i) log P(C_i) + P(t) \sum_{i=1}^{C} P(C_i|t) log P(C_i|t) + P(\bar{t}) \sum_{i=1}^{C} P(C_i|\bar{t}) log\, P(C_i|\bar{t}) \quad (8)$$

Where attribute class is represented by C_i, probability of i[th] class value is represented by $P(C_i)$, feature t occurrence probability is given by $P(t)$ and probability that feature t does not occur is given by $P(\bar{t})$, conditional probability of class that belongs to C_i is represented by $P(C_i|t)$, conditional probability that class belongs to C_i when feature t is not included is given by $P(C_i|\bar{t})$. The attribute reduction process is performed using Algorithm 2.

Algorithm 2: Attribute reduction algorithm

Input: F-Full set of attributes, IGR: Information Gain Ratio
Output: Optimal attributes
1. Begin
2. Read all the attributes F_i in the dataset
3. For all attributes, compute information entropy using equations (3) and (4)
4: Calculate the information gain ratio value for all attributes using the equation (7)
5: Rank the attributes according to the information gain ratio value of the attributes
6: Optimal attribute results
7. End

3.3.1. Apriori Algorithm to Extract Frequent Itemsets

The mining of frequent itemsets is performed by Apriori algorithm and it is the most commonly used association rule mining algorithm [2]. Itemset support is counted by using breadth-first search technique and candidate function is generated by utilizing the downward closure property. The algorithm adopts bottom up strategy and the key generation of candidates corresponds to the extension of one item at a time [6, 19].

The candidate itemsets is generated with length *k*. It prunes the candidates which have an infrequent sub pattern. All frequent *k*-length itemsets [4] are computed by using downward closure property. The frequent itemsets are computed by scanning the database. Association rule mining algorithm operates in two steps:

1. Frequent itemset generation: Itemsets with support greater than or equal to minimum support are computed.
2. Rule generation: For each frequent itemset, rules with high confidence are generated.

The new candidate itemset is generated in the first step based on the support count and the candidate set can be defined as frequent or infrequent. For generating higher level candidate itemsets (C_i) previous level frequent itemsets k_i-*1* are joined. The infrequent candidate itemsets are filtered out in the pruning step. This ensures that every subset of a frequent itemset is also frequent. Hence, if the candidate itemset contains more infrequent itemsets, it will be subsequently removed from the further process of frequent itemset generation. This process is called pruning. In the second phase, rules are generated from the frequent itemsets using the confidence measure. To select interesting rules from the set of all possible frequent itemsets, the designed system uses various measures of significance and interests. The best-known constraints are minimum thresholds set on support and confidence.

The support of an itemset *X*, $supp(X)$ represents itemset *X's* support and is defined as the ratio of transactions in which *X* appears to the total number of transactions. It signifies the popularity of an itemset. The $supp(X)$ is given by,

$$supp(X) = \frac{Number\_of\_transactions\_in\_w\square ic\square\_X\_app}{Total\_number\_of\_transactions} \quad (9)$$

The confidence of a rule is defined by,

$$conf(X \rightarrow Y) = \frac{supp(X \cup Y)}{supp\ (X)} \quad (10)$$

Where $(X \cup Y)$ is the number of transactions in which itemsets of both *X* and *Y* appears.

Frequent itemsets in the database are identified by scanning the entire database with the minimum support threshold value. Algorithm 3 represents the Apriori candidate itemset generation method.

Algorithm 3: Apriori Candidate Itemset Generation

Input: S, support where S = reduced feature set, min_support = real
Output: Set of frequent itemsets
Requirement: $S \neq \emptyset, 0 \leq min\_support \leq 1$
1. Procedure Getfrequentitemsets
2. $freqSets[\,] \leftarrow null$
3. For all itemsets i in reduced attribute set do
4. If $support \geq min\_support$ then
5. $freqSets[\,] \leftarrow i$ (frequent attribute)
6. Candidate pruning is performed to eliminate some of the candidate k-itemsets (infrequent attribute set)
7. End if
8. End for
9. End procedure

3.4. Support Vector Machine with Information Gain Based Classification

Credit card fraud is detected using SVM with information gain as given in Algorithm 4. SVMIG is an algorithm which separates the given dataset into

different classes using a hyperplane. The SVMIG determines the optimal hyperplane and the support vectors are the points closest to the hyperplane in the different classes and they are used to predict the classes of new data points. A new incoming point is classified as to which class it belongs by putting on the equation of the hyperplane on the basis of which side of hyperplane it falls on the vector space. The supervised data is given as input to the system and this is the data with results already known. The system SVMIG learns the behavior of fraud and genuine transactions and it then classifies the new transaction as to which it belongs. Information Gain based attribute selection is used to normalize the attributes and to select the better features.

Algorithm 4: Support Vector Machine with Information Gain based Classification
Testing process
Input: Test instances x_i
Output: y is the predicted class label for test instance x
1. Begin
2. For i = 1: N
3. Put x into the classifier to predict its class label y_i
4: Hyperplane separation based on weight factor
5: Credit card fraud classification
6: End for
7. End

4. Results and Discussion

Experimental results of the proposed SVMIG algorithm are evaluated using standard dataset. The experimentation is carried out on the "default of credit card clients dataset" [26] available in UCI machine learning repository and evaluated with the metrics pruning time, accuracy, precision, recall, and F1 measure.

4.1. Dataset Description

The default of credit card clients dataset contains 30000 instances and 24 attributes. The dataset represents the default payments of customers and used as an input. The dataset contains 23 variables named X1-X23 representing the explanatory variables and employs a binary variable-default payment (Yes=1, No =0), as the response variable. Variables include credit amount, gender, education, marital status, age, past payment history, bill statement, and previous payment amount.

4.2. Performance Metrics

The metrics used for evaluation here are accuracy, precision, recall, and F1 measure which are defined as given below:

- *Accuracy*: The accuracy measure of a classifier determines how uniquely the training tuple is classified by the classifier. The objective of this measure is to predict the class label of tuples and the accuracy of the classifier is estimated by its testing set. Ratio of correctly predicted observation to total observations produces accuracy value. It is the most intuitive measure of performance. Accuracy is calculated by using the following formula.

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{P+N}} \quad (11)$$

Where *TP*, *TN*, *P* and *N* refers to the number of true positive, true negative, positive, and negative samples respectively.

- *Precision*: Precision is a measure of exactness that determines what percentage of tuples is actually labeled as positive. Ratio of correctly predicted positive observations to total predicted positive observations produces precision value.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (12)$$

Where *FP* denotes the number of false positive samples.

- *Recall*: Recall is a measure of completeness that estimates what percentage of positive tuples is actually labeled. Recall is the same as sensitivity which is the true positive rate. Ratio of correctly predicted positive observations to all observations produces the value of recall in actual class. Recall is expressed as follows:

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (13)$$

Where *FN* represents the number of false negative samples.

- *F1 measure*: *F1* measure is used to determine the accuracy of the test and also it considers both recall and precision of the test to compute the score. *F1* measure specifies how precise a classifier is i.e., how many instances it correctly classifies and how robust it is i.e., it does not miss the significant number of instances. Weighted average of recall and precision produces *F1* measure. Classifier performance is rated statistically by using this measure and in this measurement, false negatives and positives are considered.

$$F1\ measure = 2\frac{Precision*recall}{Precision+recall} \quad (14)$$

4.3. Performance Analysis

The performance of the proposed SVMIG algorithm is compared with the existing methods Naive Bayesian Classifier (NBC) [9], Random-Tree-Based Random Forest (RTBRF), and CART-based Random Forest (CARTRF) based classification techniques [30].

4.3.1. Pruning Time

The pruning time of proposed SVMIG classifier and NBC, RTBRF, and CARTRF based classification

approaches are performed on the default of credit card clients dataset. Figure 2 shows the comparative representation of SVMIG against the existing methods. The results of experimentation illustrates that proposed system attains 0.0528 sec whereas other methods such as NBC, RTBRF, and CARTRF achieves 0.1633 sec, 0.1443 sec, and 0.1311 sec respectively for 1800 transactions. When the sample size is 9000, the SVMIG classifier has attained the less pruning time of 0.5365 sec since only frequent attribute sets are selected and the candidate's itemset size is reduced by Apriori algorithm. Hence, the SVMIG algorithm has consumed less time compared to the other methods.

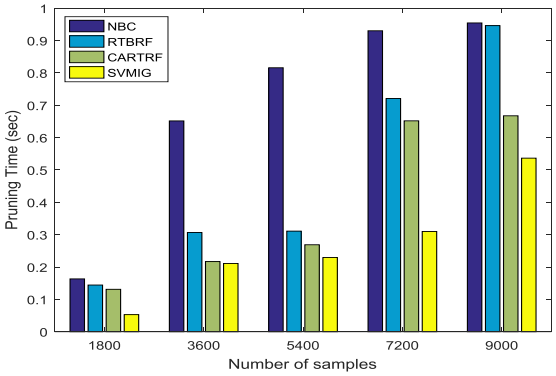

Figure 2. Pruning time estimation.

4.3.2. Accuracy

The accuracy of the proposed SVMIG algorithm and the existing methods are represented in Figure 3 which shows the accuracy comparison for default of credit card clients dataset. When the sample size is 9000, SVMIG has obtained higher accuracy value of 94.1025% while NBC, RTBRF, and CARTRF obtained accuracy values of 84.2878%, 86.8137%, and 89.6751% respectively. This is because of the use of holdout method where the dataset is randomly partitioned into training set and testing set and hence, the accuracy value of the proposed SVMIG classifier is high when the sample size is increased. Also, the robustness of the model helps in attaining enhanced classification accuracy and the good prediction of the class label.

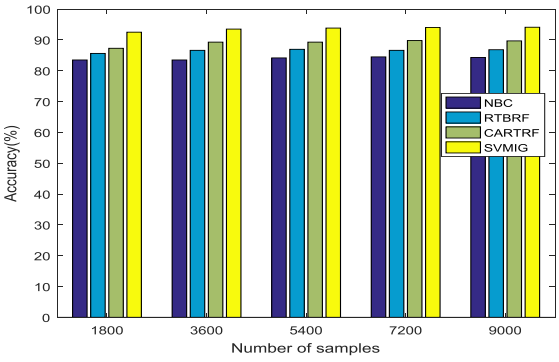

Figure 3. Accuracy measure.

4.3.3. Precision

The precision of the proposed and existing methods is represented in Figure 4 and shows the precision comparison for default of credit card clients dataset among the existing methods. When the sample size is 7200, SVMIG has obtained the precision value of 93.6055% whereas NBC, RTBRF, and CARTRF have achieved values of 83.1208%, 87.6137%, and 93.1121% respectively. The SVMIG has the highest precision value of 95.2785% when the number of sample size of the testing set is 9000 because of the exactness of classifying the legal transactions as positive and fraudulent transactions as negative. The precision value is impressive for a larger sample size because of the exact classification of legal and fraudulent transactions.

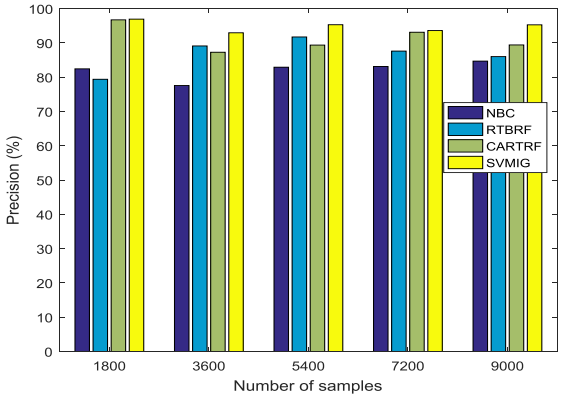

Figure 4. Precision calculation.

4.3.4. Recall

Figure 5 shows the recall comparison of different methods against the proposed method SVMIG. It is evident that when the sample size is 9000, SVMIG has obtained the highest recall value of 93.8845% while NBC, RTBRF, and CARTRF achieved recall values of 85.3338%, 90.4197%, and 88.6701% respectively. The recall value of the proposed SVMIG is high since the true positive rate that is the proportion of legal transactions is accurately determined by the classifier.

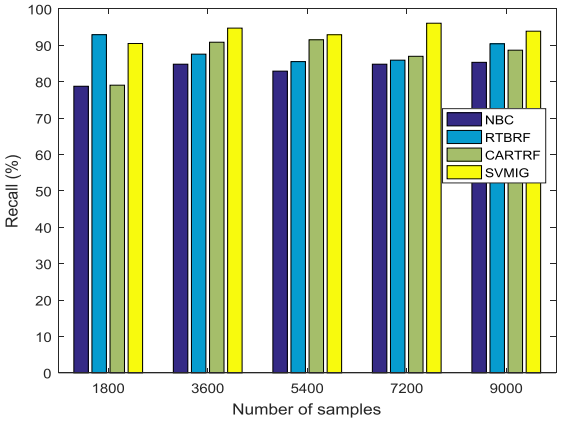

Figure 5. Recall measure.

4.3.5. F1 Measure

The F1 measure is a combined metric of both precision and recall and its comparison is performed for the default of credit card clients dataset and is represented in Figure 6. The proposed SVMIG method achieves 93.6055% of F1 measure value while other methods NBC, RTBRF, and CARTRF attains 80.5578%, 85.6137%, and 86.9891% respectively when the sample size is 1800. The SVMIG attains 94.5765% highest value whereas NBC, RTBRF, and CARTRF achieve the values of 85.0128%, 88.1587%, and 89.0391% respectively when the sample size is 9000. The proposed SVMIG attains the highest F1 score because of the accuracy of the proposed classifier.

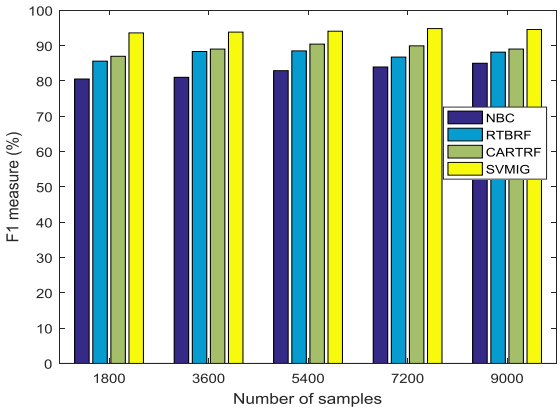

Figure 6. F1 measure calculation.

5. Conclusions and Future Work

The proposed system SVMIG identifies the fraudulent and legal transactions based on the information gain based classification. The information gain is used to select optimal attributes for the reduction of feature set. The accuracy of detection is enhanced in the credit card fraud detection system by using SVM with information gain based classification. In SVMIG, the min-max normalization is used to normalize the attribute values from dynamic range to specific range. The credit card fraud is detected using SVMIG after the selection of frequent itemsets. The detection rate is high in the proposed method for a large sample size. The standard metrics were used to evaluate the proposed algorithm and the performance of proposed SVMIG based classification has better performance against the existing methods. The SVMIG algorithm speeds up the detection convergence and attained the classification accuracy of 94.1025% when the sample size is large. The accuracy of detecting the fraudulent transaction from the legitimate ones can further be improved by using Dragonfly algorithm and Firefly algorithm to optimize the attribute reduction as a future work. The dataset available on day to day processing may become obsolete and a system is necessary for effective identification of fraud behavior in streaming data. Hence, optimization algorithms like artificial immune system, genetic algorithm, and case-based reasoning can be combined with machine learning algorithms like neural network to enhance the accuracy.

Acknowledgement

K. Poongodi gratefully acknowledges the support provided by Centre for Research, Anna University, Chennai to carry out this research work by granting Anna Centenary Research Fellowship (ACRF 2016-2018).

References

[1] Adewumi A. and Akinyelu A., "A Survey of Machine-Learning and Nature-inspired based Credit Card Fraud Detection Techniques," *International Journal of System Assurance Engineering and Management*, vol. 8, no. 2, pp. 937-953, 2017.

[2] Agrawal R. and Srikant R., "Fast Algorithms for Mining Association Rules," *in Proceedings of 20th International Conference on Very Large Databases*, San Francisco, pp. 487-499, 1994.

[3] Behera T. and Panigrahi S., "Credit Card Fraud Detection: A Hybrid Approach Using Fuzzy Clustering and Neural Network," *in Proceedings of 2nd IEEE International Conference on Advances in Computing and Communication Engineering*, Dehradun, pp. 494-499, 2015.

[4] Cai S., Hao S., Sun R., and Wu G., "Mining Recent Maximal Frequent Itemsets over Data Streams with Sliding Window," *The International Arab Journal of Information Technology*, vol. 16, no. 6, pp. 961-969, 2019.

[5] Darwish S., "An Intelligent Credit Card Fraud Detection Approach based on Semantic Fusion of Two Classifiers," *Soft Computing*, vol. 24, no. 2, pp. 1243-1253, 2020.

[6] Datta D., De A., Roy D., and Dutta S., "Apriori Algorithm using Hashing for Frequent Itemsets Mining," *International Journal of Education and Management Engineering*, vol. 8, no. 6, pp. 46-58, 2018.

[7] De Sá A., Pereira A., and Pappa G., "A Customized Classification Algorithm for Credit Card Fraud Detection," *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 21-29, 2018.

[8] Delamaire L., Abdou H., and Pointon J., "Credit Card Fraud and Detection Techniques: a Review," *Banks and Bank Systems*, vol. 4, no. 2, pp. 57-68, 2009.

[9] Domingos P. and Pazzani M., "Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier," *in Proceedings*

of the 13th International Conference on Machine Learning, Bari, pp. 105-112, 1996.

[10] Duman E. and Ozcelik M., "Detecting Credit Card Fraud by Genetic Algorithm and Scatter Search," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13057-13063, 2011.

[11] Excell D., "Bayesian Inference-The Future of Online Fraud Protection," *Computer Fraud and Security*, vol. 2012, no. 2, pp. 8-11, 2012.

[12] Fu K., Cheng D., Tu Y., and Zhang L., "Credit Card Fraud Detection using Convolutional Neural Networks," *in Proceedings of the International Conference on Neural Information Processing*, Kyoto, pp. 483-490, 2016.

[13] Jain Y., Tiwari N., Dubey S., and Jain S., "A Comparative Analysis of Various Credit Card Fraud Detection Techniques," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5S2, pp. 402-407, 2019.

[14] Jha S., Guillen M., and Westland J., "Employing Transaction Aggregation Strategy to Detect Credit Card Fraud," *Expert Systems with Applications*, vol. 39, no. 16, pp. 12650-12657, 2012.

[15] Jiang C., Song J., Liu G., Zheng L., and Luan W., "Credit Card Fraud Detection: A Novel Approach using Aggregation Strategy and Feedback Mechanism," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3637-3647, 2018.

[16] Khare N. and Sait S., "Credit Card Fraud Detection using Machine Learning Models and Collating Machine Learning Models," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 20, pp. 825-838, 2018.

[17] Maes S., Tuyls K., Vanschoenwinkel B., and Manderick B., "Credit Card Fraud Detection using Bayesian and Neural Networks," *in Proceedings of the 1st International Naiso Congress on Neuro Fuzzy Technologies*, Havana, Cuba, pp. 261-270, 2002.

[18] Ogwueleka F., "Data Mining Application in Credit Card Fraud Detection System," *Journal of Engineering Science and Technology*, vol. 6, no. 3, pp. 311-322, 2011.

[19] Pei J., Han J., Mortazavi-Asl B., Wang J., Pinto H., Chen Q., Dayal U., and Hsu M., "Mining Sequential Patterns by Pattern-growth: The Prefixspan Approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 11, pp. 1424-1440, 2004.

[20] Quah J. and Sriganesh M., "Real-time Credit Card Fraud Detection using Computational Intelligence," *Expert Systems with Applications*, vol. 35. no. 4, pp. 1721-1732, 2008.

[21] Randhawa K., Loo C., Seera M., Lim C., and Nandi A., "Credit Card Fraud Detection using AdaBoost and Majority Voting," *IEEE Access*, vol. 6, pp. 14277-14284, 2018.

[22] Razooqi T., Khurana P., Raahemifar K., and Abhari A., "Credit Card Fraud Detection using Fuzzy Logic and Neural Network," *in Proceedings of 19th Communications and Networking Symposium*, Pasadena, pp. 1-5, 2016.

[23] Srivastava A., Kundu A., Sural S., and Majumdar A., "Credit Card Fraud Detection using Hidden Markov Model," *IEEE Transactions on Dependable and Secure Computing*, vol. 5, no. 1, pp. 37-48, 2008.

[24] Teh B., Islam M. B., Kumar N., Islam M., and Eaganathan U., "Statistical and Spending Behavior based Fraud Detection of Card-based Payment System," *in Proceedings of the International Conference on Electrical Engineering and Informatics*, Banda Aceh, pp. 78-83, 2018.

[25] Tripathi K. and Pavaskar M., "Survey on Credit Card Fraud Detection Methods," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 11, pp. 721-726, 2012.

[26] UCI Machine Learning Repository, default of credit card clients dataset, https://archive.ics.uci.edu/ml/datasets/default+of +credit+card+clients, Last Visited, 2020.

[27] Van Vlasselaer V., Bravo C., Caelen O., Eliassi-Rad T., Akoglu L., Snoeck M., and Baesens B., "APATE: A Novel Approach for Automated Credit Card Transaction Fraud Detection using Network-based Extensions," *Decision Support Systems*, vol. 75, pp. 38-48, 2015.

[28] Wang C. and Han D., "Credit Card Fraud Forecasting Model based on Clustering Analysis and Integrated Support Vector Machine," *Cluster Computing*, vol. 22, no. 6, pp. 13861-13866, 2019.

[29] Wang Y., Adams S., Beling P., Greenspan S., Rajagopalan S., Velez-Rojas M., Mankovski S., Boker S., and Brown D., "Privacy Preserving Distributed Deep Learning and its Application in Credit Card Fraud Detection," *in Proceedings of 17th IEEE International Conference on Trust, Security and Privacy*, New York, pp. 1070-1078, 2018.

[30] Xuan S., Liu G., Li Z., Zheng L., Wang S., and Jiang C., "Random Forest for Credit Card Fraud Detection," *in Proceedings of 15th IEEE International Conference on Networking, Sensing and Control*, Zhuhai, pp. 1-6, 2018.

[31] Zaslavsky V. and Strizhak A., "Credit Card Fraud Detection using Self-Organizing Maps," *Information and Security*, vol. 18, pp. 48-63, 2006.

Kannan Poongodi is a Ph.D. student in the Department of Information Technology, Anna University, MIT Campus, Chennai. She received the MCA degree with distinction from Bharathiar University, Coimbatore in 2003. She was awarded M.E. degree in Computer Science and Engineering with distinction and received 39th rank from Anna University, Chennai in 2014. She has nine years of teaching experience and received best teacher award in the year 2012 by Akshaya Institute of Management Studies, Coimbatore in the Coimbatore District. Her research interests include Data Mining, Big Data, Database Management System and Computer Networks.

Dhananjay Kumar received his bachelor degree in Electronics and Telecommunication engineering from the Institution of Engineers (India), Calcutta in the year 1997. He was awarded a Master of Engineering (M.E.) degree in Industrial Electronic Engineering by the Maharaja Sayajirao University of Baroda in the year 1999. He also holds a Master of Technology (M.Tech.) degree in Communication Engineering from Pondicherry University through Pondicherry Engineering College, Pondicherry in the year 2001. A Doctor of Philosophy (Ph.D.) degree was awarded to him by Anna University, Chennai in the year 2009 for his research work on high quality multimedia support in next generation multi-carrier wireless mobile networks. Currently, he is working as Professor and Head in the Department of Information Technology, Anna University, MIT Campus, Chennai. Dr. Kumar and his team are developing a machine learning based system to support high-quality video streaming over internet through future wireless networks. Additionally, a group of researchers are working under his guidance for the detection and tracking of object in the live stream of video and data analysis.