**VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI, KARNATAKA, INDIA**

****

Text

Description automatically generated with medium confidence

**An Autonomous Institution with A+ Grade UGC by NAAC UGC, Approved by UGC, AICTE, Government of Karnataka, Yelahanka, Bengaluru-560064, Karnatka, India.**

**Mini Project Report On**

**“DATA CLASSIFICATION USING SVM”**

A Mini Project report submitted in partial fulfillment of the requirement for the award of

**BACHELOR OF ENGINEERING**

**IN**

**ELECTRONICS AND COMMUNICATION ENGINEERING**

**2021-2022**

Submitted By

**Student Name USN**

Kushal S 1NT18EC081

Brunda S A 1NT18EC037

**Under the Guidance of**

Dr. H.S. Prashanth

Professor

Dept. of Electronics and Communication Engineering

Nitte Meenakshi Institute of Technology

Yelahanka, Banglore-64

Text

Description automatically generated with medium confidence

**An Autonomous Institution with A+ Grade UGC by NAAC UGC, Approved by UGC, AICTE, Government of Karnataka, Yelahanka, Bengaluru-560064, Karnatka, India.**

**Department of Electronics and Communication Engineering**

Certificate

*This is to certify that* ***Brunda S A****(****1NT18EC037****) and* ***Kushal S (1NT18EC081)*** *has submitted the Mini Project report entitled* ***“DATA CLASSIFICATION USING SVM”*** *in fulfillment for the award of Bachelor of Engineering**in**Electronics and Communication Engineering**from**Visvesvaraya Technological University, Belagavi during the year 2021-2022. It is certified that all the corrections, suggestions indicated for internal assessment have been incorporated in the report. The Mini Project report has been approved as it satisfies the academic requirements in respect of work prescribed for the aforesaid degree.*

Dr. H.S Prashanth Dr. Ramachandra A C Dr. H. C. Nagaraj

**Guide HOD**  **Principal**

**External Viva Voce**

**Name of Examiners Signature with Date**

**1) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**2) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

***ACKNOWLEDGEMENT***

It is my proud privilege and duty to acknowledge the kind help and guidance received from several people in preparation of this report. It would not have been possible to prepare this report in this form without valuable suggestions, cooperation and guidance.

I wish to record my sincere gratitude to Management, and **Dr. H. C. Nagaraj,** Principal, Nitte Meenakshi Institute of Technology, Bengaluru for the permission provided to accomplish this Mini Project.

My sincere thanks to **Dr. Ramachandra A. C,** Professor and Head, Department of Electronics and Communication Engineering for his valuable suggestions and guidance.

My sincere gratitude to **Dr. Rajesh N, Dr. Sunil S. Harakannanavar** and **Dr. Thimmaraja Yadava G,** Mini Project Coordinators**,** Department of Electronics and Communication Engineering, for their valuable suggestions in preparing this report.

I express my sincere gratitude to our beloved guide**, Dr. H.S Prashanth,** Professor, Department of Electronics and Communication and Engineering for his support and guidance.

I am extremely great full to faculty members of Department of Electronics and Communication Engineering, and friends for their support and encouragement in successful completion of this Mini Project.

**Name of Student USN**

Brunda S A 1NT18EC037

Kushal S 1NT18EC081

**Contents**

1. Abstract Page No.7
2. Introduction Page No.8-10
3. Literature Survey Page No.11-13
4. Objectives of the Project Page No.14
5. Design Methodology/block diagram/flowchart Page No.14
6. Implementation/technical description of the project Page No.15
7. Experimental results and analysis Page No.26
8. Conclusions Page No.27

References Page No.28

**List of Figures**

Figure 1: Block diagram of how payment process works Page No.9

Figure2: Design Methodology Page No.14

Figure3: Data Visualization Page No.15

Figure 4: ROC curve of Naive Bayes Page No.17

Figure 5: Linear regression graph Page No.18

Figure 6: Sigmoid function Page No.19

Figure 7: confusion matrix Page No.19

Figure 8: Logistic Regression Degree 1 Page No.20

figure 9: Logistic Regression Degree 2 Page No.21

Figure 10: Support vector Machine Page No.23

Figure 11: SVM Kernel rbf Page No.24

Figure 12: SVM Sigmoid Page No.24

Figure 13: SVM Kernel Linear Page No.25

**List of Tables**

Table 1: Comparison of results of different classification algorithm Page No.26

**Abstract**

Financial fraud is an ever-growing menace with far consequences in the financial industry. Data mining has played an imperative role in the detection of credit card fraud in online transactions. Credit card fraud detection, which is a data mining problem, becomes challenging due to two major reasons - first, the profiles of normal and fraudulent behaviour change constantly and secondly, credit card fraud data sets are highly skewed.

The performance of fraud detection in credit card transactions is greatly affected by the sampling approach on the dataset, selection of variables and detection technique(s) used. This paper investigates the performance of naïve bayes, logistic regression and many others on highly skewed credit card fraud data detection.

A hybrid technique of under-sampling and oversampling is carried out on the skewed data. The three techniques are applied on the raw and preprocessed data. The work is implemented in Python. The performance of the techniques is evaluated based on accuracy, sensitivity, specificity, precision.

The results show the optimal accuracy for naïve bayes, logistic regression and support vector machine algorithms respectively.The comparative results show that a support vector machine performs better than naive bayes and logistic regression.

**Chapter 1**

**Introduction**

**Machine learning** was defined in the 90's by **Arthur Samuel** described as the,” it is a field of study that gives the ability to the computer for self-learning without being explicitly programmed”, that means imbuing knowledge to machines without hard-coding it.

* It can process massive data faster with the learning algorithm. For instance, it will be interested in learning to complete a task, make accurate predictions, or behave intelligently.
* Why Machine Learning?

To develop systems that can automatically adapt and customize themselves to individual users. It can mimic human and replace certain monotonous tasks, which require some intelligence from large data bases (Data mining). We can also implement it on difficult/expensive to construct to manually because they require specific detailed skills and knowledge tuned to a specific task.

**SUPPORT VECTOR MACHINES (SVM**)**:**

In machine learning, **Support vector machines (SVM**) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. It is mostly used in classification problems. In this algorithm, each data item is plotted as a point in n-dimensional space (where n is the number of features), with the value of each feature being the value of a particular coordinate. Then, classification is performed by finding the hyper-plane that best differentiates the two classes.

**CREDIT CARD**:

Credit cards offer you a line of credit that can be used to make purchases, balance transfers or cash advances and require that you pay back the loan amount in the future.

* When using a credit card, you will need to make at least the minimum payment every month by the due date on the balance.

#### What is credit card fraud detection?

It’s a set of activities aimed to save the money and assets being stolen by criminals. There are a lot of traditional methods, which worked perfectly fine for some time.

Chart

Description automatically generated

However now, it is a kind of area where constant improvement is crucial for survival. That's why it is important to talk about what Machine Learning, a subdivision of Artificial Intelligence, can do here. You need a team of Data Science experts and main information on credit card transactions in your organization, such as Amount, Date, User Location, Usual Behavioral Patterns of your clients, etc. Experts, obtaining that, will be able to develop a model that will be able to prevent and reveal any fraudulent activity among your transactions. This model will be fine-tuned in the process, and there are few ways to make it work, we will talk about it later. Now, let's move on to the most common scenarios, which ML powered Fraud Detection could handle easily.

#### **Causes of credit card fraud:**

* This growth in losses is partially caused by the rise of electronic transactions.
* Diagram

  Description automatically generatedFraudulent methods are getting more sophisticated and thus harder to spot by traditional fraud detection methods.

**Fig1: block diagram of how payment process works**

All players involved in the card-based payment process can potentially fall victim to scammers, including:

* cardholders,
* online merchants,
* payment gateway providers,
* payment processing companies,
* credit card payment systems,
* card issuers (issuing banks), and
* acquirers (acquiring banks).

**How machine learning helps with fraud detection:** The key objective of any credit card fraud detection system is to identify suspicious events and report them to an analyst while letting normal transactions be automatically processed. But now they increasingly turn to a machine learning approach, as it сan bring significant improvements to the process.

#### **Higher accuracy of fraud detection:**

Compared to rule-based solutions, machine learning tools have higher precision and return more relevant results as they consider multiple additional factors. This is because ML technologies can consider many more data points, including the tiniest details of behavior patterns associated with a particular account.

#### **Less manual work needed for additional verification:**

Enhanced accuracy leads reduce the burden on analysts. People are unable to check all the transactions manually even in small banks.

#### **Fewer false declines:**

False declines or false positives happen when a system identifies a legitimate transaction as suspicious and wrongly cancels it. unsupervised learning techniques may be applied to analyze available data and find patterns of fraudulent behavior.

#### **Ability to identify new patterns and adapt to changes:**

Unlike rule-based systems, ML algorithms are aligned with a constantly changing environment and financial conditions. They enable analysts to identify new suspicious patterns and create new rules to prevent new types of scams

**Chapter 2**

**Literature Survey**

**STUDY ON SUPPORT VECTOR MACHINE BASED LINEAR AND NON-LINEAR PATTERN CLASSIFICATION:**

In this paper concept of SVM along with its various models has been presented. Phenomenon which leads to the concept of SVM back in the 90s’ is mentioned. SVMs are mainly used for classification and regression analysis both for linear and nonlinear decision boundaries.

**COMPARATIVE STUDY OF BIG DATA CLASSIFICATION ALGORITHM BASED ON SVM:**

The emergence of linear support vector machines algorithm package LIBLINEAR has solved the problem. In fact, LIBLINEAR is originally designed to solve the problem of large amount of data. Using completely different optimization algorithms from LIBSVM, LIBLINEAR greatly reduces training computational complexity and time consumption while maintaining similar effects in linear SVM classification**.**

**SUPPORT VECTOR MACHINE ACCURACY IMPROVEMENT WITH CLASSIFICATION:**

The report represents superiority to the presented approach (SVM based on RBF) which delivers extensively enhanced value against the other classifiers. The misclassification value is removed from the final Support Vector Classification. An improved accuracy is obtained after a certain modification on the epsilon and cost parameter in radial basis function. It has been observed in the experiment that the trained SVM gives 98% accuracy.

**CLASSIFICATION MECHANISM OF SUPPORT VECTOR MACHINES:**

This formulation results in a global quadratic optimization problem with convex constraints, which is readily solved by interior point methods. The kernel mapping provides a unifying framework for most of the commonly employed model architectures.

**RESEARCH ON SVM IMPROVED ALGORITHM FOR LARGE DATA CLASSIFICATION:**

In view of the two problems of the SVM algorithm in processing large data, the paper proposed a weighted Euclidean distance, radial integral kernel function SVM and dimensionality reduction algorithm for large data packet classification. The improved algorithm reconstructs the data feature space, makes the boundary of different data samples clearer, shortens the modeling time, and improves the accuracy of classification.

**IMPROVING CLASSIFICATION ACCURACY USING SVMS WITH NEW KERNEL:**

In this paper, they introduce a new kernel function called polynomial radial basis function (PRBF) that could improve the classification accuracy of support vector machines (SVMs). The proposed kernel function combines both Gauss (RBF) and Polynomial (POLY) kernels and is stated in general form. It is shown that the proposed kernel converges faster than the Gauss and Polynomial kernels.

**BEHAVIOR BASED CREDIT CARD FRAUD DETECTION USING SUPPORT VECTOR MACHINES**

Along with the great increase of internet and e-commerce, the use of credit card is an unavoidable one. Due to the increase of credit card usage, the frauds associated with this have also increased. There are a lot of approaches used to detect the frauds. In this paper, behavior-based classification approach using Support Vector Machines are employed and efficient feature extraction method also adopted. If any discrepancies occur in the behavior’s transaction pattern, then it is predicted as suspicious and taken for further consideration to find the frauds. Generally, credit card fraud detection problem suffers from a large amount of data, which is rectified by the proposed method. Achieving finest accuracy, high fraud catching rate and low false alarms are the main tasks of this approach.

**IMPLEMENTATION OF CREDIT CARD FRAUD DETECTION USING SUPPORT VECTOR MACHINE**

Credit card fraud has without hesitation an expression of criminal deception. Fraud identification seems to be a complicated problem that requires a significant amount of skill until throwing algorithms regarding machine learning into it. However, it is an implementation for both the better of machine learning as well as artificial intelligence, ensuring that perhaps the funds of both the customer seems to be secure and therefore not manipulated. The whole research article addressed an effective system of identifying fraud depending on machine learning methodologies, with such a feedback system. Its feedback process relates to enhancing the classifier's detection rate as well as effectiveness. The state of art approaches failed to detect the frauds through credit card transactions. Thus, to solve these drawbacks, the proposed system is implemented with the Support vector machine (SVM) classification to detect the frauds. The simulation results show that the proposed method gives the better classification accuracy compared to the state of art approaches.

**CREDIT CARD FRAUD DETECTION USING WEIGHTED SUPPORT VECTOR MACHINE**

Credit card fraudulent data is highly imbalanced, and it has presented an overwhelmingly large portion of nonfraudulent transactions and a small portion of fraudulent transactions. The measures used to judge the veracity of the detection algorithms become critical to the deployment of a model that accurately scores fraudulent transactions taking into account case imbalance, and the cost of identifying a case as genuine when, in fact, the case is a fraudulent transaction. In this paper, a new criterion to judge classification algorithms, which considers the cost of misclassification, is proposed, and several undersampling techniques are compared by this new criterion. At the same time, a weighted support vector machine (SVM) algorithm considering the financial cost of misclassification is introduced, proving to be more practical for credit card fraud detection than traditional methodologies. This weighted SVM uses transaction balances as weights for fraudulent transactions, and a uniformed weight for nonfraudulent transactions. The results show this strategy greatly improve performance of credit card fraud detection.

**SELECTION FEATURES AND SUPPORT VECTOR MACHINE FOR CREDIT CARD RISK IDENTIFICATION**

For identifying credit card risk in massive and high dimensionality data, feature selection is considered very important to improve classification performance and fraud identification process. One of the commonly used feature selection methods is Random Forest Classifier (RFC), which is very suitable for large dataset. RFC has a good performance; it tends to identify the most predictive features, which may provide a significant improvement in classification performance of credit card risk identification model. In this paper, we propose an enhanced Credit Card Risk Identification (CCRI) method based on the features selection algorithm as Random Forest Classifier and Support Vector Machine to detecting fraud risk. Our experimental results show that the proposed algorithm outperforms the Local Outlier Factor, Isolation Forest and Decision Tree in term of classification performance on a larger dataset.

**Chapter 3**

**Project Objectives**

The following are the proposed objectives

* To develop a credit card fraud detection mechanism using machine learning.
* To implement SVM, Naive Bayes, Logistic Regression method for data classification.
* To find the accuracy of all the algorithms used.
* To compare different algorithms and find best suited algorithm for credit card fraud detection.

**Chapter 4**

**Diagram

Description automatically generatedSystem Methodology/Block Diagram/Proposed Solution/Design methodology**

**Fig2: block diagram**

**Chapter 5**

**Implementation/Technical Description of the Project**

**Data Visualization**

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets.

The term is often used interchangeably with others, including information graphics, information visualization and statistical graphics.

Graphical user interface, application

Description automatically generated**Fig3: Data Visualization**

Chart, bar chart

Description automatically generated

## Data Binning:

* Binning is a way to group a number of more or less continuous values into a smaller number of "bins".
* For example, if you have data about a group of people, you might want to arrange their ages into a smaller number of age intervals.
* Numeric columns can also be temporarily grouped by right-clicking on a column selector and clicking Auto-bin Column.

Chart, waterfall chart

Description automatically generated

## Implementation of Algorithms for Credit Card Fraud Detection.

1. **Support Vector Machine**
2. **Naive Bayes**
3. **Logistic regression.**

## Naive Bayes:

Naive Bayes is one of the most popular and simple machine learning classification algorithms, in Naive Bayes algorithm the classification technique based on Bayes Theorem with an assumption of independence among predictors. In simple terms, the classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. It’s easy to build and particularly useful for large data sets, hence its known to outperform even highly sophisticated classification models and can be implemented easily.

#### **Using Probabilistic Model**:

Text, letter

Description automatically generated with medium confidenceWhen we try to calculate probability on a condition, i.e., probability of happening of event A when event B has already taken place.

* **Likelihood:** Indicates how probable is the evidence given that our hypothesis is true.
* **Posterior Probability:** Indicates how probable is our hypothesis given from the observed evidence (Not directly Computable).
* **Predictor Prior Probability:** Indicates how probable is the new evidence under all possible hypotheses.
* **Class Prior Probability:** Indicates how probable our Hypotheses are before observing the evidence.

#### **Steps to implement Naive Bayes Algorithm:**

**Step 1:** Handling the data and summarizing. It can even involve converting the table into frequency.

**Step 2:** Making predictions and creating likelihood to evaluate accuracy.

**Step 3:** Trying all together and to calculate the posterior probability of each class. The class with the highest prediction is considered for the ideal output.

#### **Applications of Naive Bayes Algorithms:**

1. Text classification/ Spam Filtering/ Sentiment Analysis
2. Recommendation System: Naive Bayes Classifier uses machine learning and data mining techniques to filter unseen information and predict whether a user would like a given resource or not.

**Output/Graph:**

**Fig 4: ROC curve of Naive Bayes**

Chart, line chart

Description automatically generated

## Logistic Regression:

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes. In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).

Diagram

Description automatically generatedMathematically, a logistic regression model predicts P(Y=1) as a function of X. It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

**Fig5: Linear regression graph**

#### **Types of Logical Regression:**

* **Binary or Binomial:** In such a kind of classification, a dependent variable will have only two possible types either 1 and 0. For example, these variables may represent success or failure, yes or no, win or loss etc.
* **Multinomial:** In such a kind of classification, dependent variables can have 3 or more possible unordered types or the types having no quantitative significance. For example, these variables may represent “Type A” or “Type B” or “Type C”.
* **Ordinal:** In such a kind of classification, dependent variables can have 3 or more possible *ordered* types or the types having a quantitative significance. For example, these variables may represent “poor” or “good”, “very good”, “Excellent” and each category can have the scores like 0,1,2,3.

In the Credit Card Fraud Detection project, we use **Binary Type of Logical Regression.**

#### **Sigmoid Function:**

Logistic Regression uses a more complex cost function also known as Sigmoid Function. Hypothesis of logistic regression tends to limit the cost function instead of a linear function. The hypothesis of logistic regression tends to limit the cost function between 0 and 1. Therefore linear functions fail to represent it as it can have a value greater than 1 or less than 0 which is not possible as per the hypothesis of logistic regression.

**A picture containing night sky

Description automatically generatedChart

Description automatically generated with medium confidenceFig6: Sigmoid function**

#### **Confusion matrix:**

A confusion matrix summarizes the classification performance of a classifier with respect to some test data. It is a two-dimensional matrix, indexed in one dimension by the true class of an object and in the other by the class that the classifier assigns. Implemented for predictive learning.

**Fig7: confusion matrix**

Graphical user interface, application

Description automatically generated

#### **Steps To implement Logistic Regression:**

To implement the Logistic Regression, we will use the same steps as we have done in previous topics of Regression.

**step 1:** Data Preprocessing, we will pre prepare the data so that we can use it in our code.

**step 2:** Fitting Logistic Regression to the Training set.

**Step 3:** Predicting the test result.

**Step 4:** Test accuracy of the result (Creation of Confusion matrix) and the verification process is done at this step.

#### **Application of Logistic Regression:**

1. Sentiment Analysis: Analyzing the sentiment using the review or tweets is the objective of this use case. Most of the brands and companies use this to increase customer experience.
2. Image segmentation, recognition and classification: The objective of all these use cases is to identify the object in the image and classify it.

Chart, line chart

Description automatically generated**fig 8: Logistic Regression Degree 1**

## Output/Graph

**fig 9: Logistic Regression Degree 2**

Chart, line chart

Description automatically generated

## Support Vector Machine Algorithm:

Support Vector Machine (SVM) it is a supervised machine learning algorithm used for both classification and regression and it is particularly applicable for small and medium sized datasets. Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in future.

Diagram

Description automatically generated

This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called support vectors, and hence the algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane.

Diagram, schematic

Description automatically generated

#### **Types of SVM:**

1. **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.
2. **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

**Hyperplane and Support Vectors in the SVM algorithm:**

**Hyperplane**:

There can be multiple lines/decision boundaries to segregate the classes in n-dimensional space, but we need to find out the best decision boundary that helps to classify the data points. This best boundary is known as the hyperplane of SVM.The dimensions of the hyperplane depend on the features present in the dataset, which means if there are 2 features (as shown inimage), then the hyperplane will be a straight line. And if there are 3 features, then the hyperplane will be a 2-dimension plane. We always create a hyperplane that has a maximum margin, which means the maximum distance between the data points.

**Support Vectors: :**

The data points or vectors that are the closest to the hyperplane and which affect the position of the hyperplane are termed as Support Vector. These vectors support the hyperplane, hence called a Support vector.

#### **Applications of Support Vector Machine Algorithm:**

1. Handwriting analysis.
2. Protein fold and remote homology detection.

**Steps to implement support vector machine algorithm:**

**Step 1**: Collect a training set.

**Step2:** Choose a kernel and its parameters as well as any regularization needed.

**Step3:** Extract from the correlation matrix.

**Step 4:** Train your machine exactly or approximately to get contraction coefficients.

**Step 5:** Use the coefficients to create your estimator.

## Output/Graphs:

## Chart, line chart Description automatically generatedFig 10: Support vector Machine

Page **25** of **29**

**Fig 11: SVM Kernel rbf**

**Chart, line chart

Description automatically generated**

**fig 12: SVM SigmoidChart, line chart

Description automatically generated**

**Chart, line chart

Description automatically generatedfig 13: SVM Kernel Linear**

**Chapter 6**

**Table

Description automatically generated with medium confidenceExperimental Results and Analysis**

Table1: Comparison of results of different classification algorithm

**Chapter 7**

**Conclusions**

By seeing the results, we can clearly conclude that Support Vector Machine (SVM) is the best performing algorithm over Logistic regression & Naïve Bayes. So, we can use SVM in production environment to predict whether the transaction is fraud or not.

|  |
| --- |
| **References**  [1]. Ghosh, S., Dasgupta, A., & Swetapadma, A. (2019). A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification. 2019 International Conference on Intelligent Sustainable Systems (ICISS). doi:10.1109/iss1.2019.8908018  [2]. Afif, M. H., & Hedar, A.-R. (2012). Data classification using Support Vector Machine integrated with scatter search method. 2012 Japan-Egypt Conference on Electronics, Communications and Computers. doi:10.1109/jec-ecc.2012.6186977  [3]. Aurangabadkar, S., & Potey, M. A. (2014). Support Vector Machine based classification system for classification of sport articles. 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). doi:10.1109/icicict.2014.6781268  [4]. Zou, H., & Jin, Z. (2018). Comparative Study of Big Data Classification Algorithm Based on SVM. 2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). doi:10.1109/csqrwc.2018.8455423  [5]. H. Dai, "Research on SVM improved algorithm for large data classification," 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), 2018, pp. 181-185, doi: 10.1109/ICBDA.2018.8367673.  [6]. H. Liu, X. Xiao, Y. Li, Q. Mi, and Z. Yang, "Effective Data Classification via Combining Neural Networks and SVM," 2019 Chinese Control and Decision Conference (CCDC), 2019, pp. 4006-4009, doi: 10.1109/CCDC.2019.8832442.  [7]. A. Gupta, N. Chetty and S. Shukla, "A classification method to classify high dimensional data," 2015 International Conference on Computing, Communication and Security (ICCCS), 2015, pp. 1-6, doi: 10.1109/CCCS.2015.7374132.  [8]. Awoyemi, J.O., Adetunmbi, A.O. and Oluwadare, S.A., 2017, October. Credit card fraud detection using machine learning techniques: A co parative analysis. In 2017 International Conference on Computing Networking and Informatics (ICCNI) (pp. 1- 9). IEEE.  [9]. Adewumi, A.O. and Akinyelu, A.A., 2017. A survey of machine-learning and natureinspired based credit card fraud detection techniques. International Journal of System Assurance Engineering and Management, 8(2), pp.937-953  [10]. Fu, K., Cheng, D., Tu, Y. and Zhang, L., 2016, October. Credit card fraud detection using convolutional neural networks. In International Conference on Neural Information Processing (pp. 483-490). Springer, Cham.  [11]. Carneiro, N., Figueira, G. and Costa, M., 2017. A data mining-based system for credit-card fraud detection in e-tail. Decision Support Systems, 95, pp.91-101.  [12]. Dhankhad, S., Mohammed, E. and Far, B., 2018, July. Supervised machine learning algorithms for credit card fraudulent transaction detection: a comparative study. In 2018 IEEE International Conference on Information Reuse and Integration (IRI) (pp. 122-125). IEEE.  [13]. Adewumi, A.O. and Akinyelu, A.A., 2017. A survey of machine-learning and natureinspired based credit card fraud detection techniques. International Journal of System Assurance Engineering and Management, 8(2), pp.937-953.  [14]. Fiore, U., De Santis, A., Perla, F., Zanetti, P. and Palmieri, F., 2019. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, pp.448-455.  [15]. Popat, R.R. and Chaudhary, J., 2018, May. A survey on credit card fraud detection using machine learning. In 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1120-1125). IEEE.  [16].] Patil, S., Nemade, V. and Soni, P.K., 2018. Predictive modelling for credit card fraud detection using data analytics. Procedia computer science, 132, pp.385-395. |