Upi Fraud Detection Using Particle Swarm Optimization Algorithm Based on the Machine Learning

Mr. Khadir Kumar.N[1], Ramesh N[2], Muralidharan P[3], Sarathi M[4], Yuvaraj[5]
[1]*HOD, Department of Computer Science and Engineering, Maha Barathi Engineering College (Affiliated to Anna University), (Affiliated to Anna University), Chinnasalem (Tk), Kallakurichi (Dt)-606 201.*
[2,3,4,5] *UG Student, Department of Computer Science and Engineering, Maha Barathi Engineering College (Affiliated to Anna University), (Affiliated to Anna University), Chinnasalem (Tk), Kallakurichi (Dt)-606 201.*

Abstract- **In the last decade, there have been several instances of fraud in the credit process; the fraud issues automatic early detection of the process. The drawback is customer dissatisfaction, increased customer service inquiries, and potentially loss of the process. We proposed method is Support Vector Machine – Decision Tree (SVM-DT) for classification and fraud detection in the process. The fraud detection of the process is more accurate of the result using Machine Learning (ML), and the analysis of the real-time dataset is used for Deep Learning (DL), a method based on Artificial Intelligence technology. The presented method is used to maintain the standard level of the reliability process. The proposed method is a classification more dataset in a particular dataset, and more accurate level of performance. The presented techniques are a breakthrough in prediction and early fraud detection prevention; it's one of the most advanced techniques in the process. Particle swarm optimization algorithm (PSO) used for the is a narrow channel of the best solution in the process, and individual personal data is enhanced a secure level, and more efficient in the stability of the process is a feature selection of the dataset. The evaluated method reduces the environment impact and time-consuming nature of the process, and the detect of misclassification and changing channels of the process is solved for the techniques. This paper's research will provide a reference for improving credit card fraud detection accuracy and efficiency.**

Keywords- **Machine Learning, Deep Learning, SVM-DT, and POS.**

I.INTRODUCTION

The Machine Learning techniques are a more advanced technology, and early detection of credit fraud, and the prediction of the result is more efficiency in the process. The ML is based on Artificial Intelligence (AI), the detection of the process is more accurate than the result, and has high reliability across various methods [1]. The presented method quickly identifies the mismatching data and selects the correct path of the network connection process. The process is an enhancement in the secure level, and collects all types of dataset information in the techniques. The Deep Learning (DL) is a real-time data analysis of the process. It compares the post and previous data, focusing on data-driven preservation and standard accuracy of the process in the Support Vector algorithm [2]. The Adaboost algorithm is used as one of the ensembles learning methods and is based on ML technology. The process of the algorithm is a weak classifier to improve the strong classifier of the process.

AI technology is most important in preventing credit fraud detection. DL performs well in analyzing all kinds of datasets and improves the network process in the KNN method. The Synthetic Minority Over-Sampling Technique (SMOTE) is a particular type of analysis that is more significant number of datasets and reduces the time complexity [3]. The ML techniques are a hybrid model for the Long Short-Term Memory (LSTM), is handles all types of data collected from the process, similar to the Gated Recurrent Unit (GRU). The process is maintained in the scalability and standard of the stability of the process. The Naive Bayes method analyzes the significant danger information, and the result is avoided, and the process can handle any critical situation with techniques. The presented method is a coding implementation process that quickly improves performance and multitasking if the tasks are managed simultaneously.

Credit fraud is one of the most critical problems in the digital technology process, affecting the organization and the whole city. The threat detection, prevention, and response method used in the AI technique is secure in the datasets and identifies the mismatching dataset [4]. The traditional signature method is ineffective in credit fraud detection in the digital technology process; the presented method collects more data but has a limited capacity range. The process creates more network connections because the builds on trustworthiness, but the process is complex because more space is occupied [5]. However, the ML development process is expensive, and the practical implementations are complex because of the environmental impacts. The DL methods are a complicated and time-consuming coding implementation process that is complicated and more time-consuming, and the process involves testing, verification, and measuring the process, which incurs high computational costs.

The main contribution of the process is that an collected the dataset for classification of the binary type of data, and data select in the hyperplane for the margin of the dataset of the network connection of the process. Decision Tree is a fraud detection process in measuring techniques and dataset classification. The proposed method is a narrow channel of the best solution in the process, and individual personal data is enhanced to a secure level, and a classification of the dataset is more efficient in the stability of the process. The process is the shortest path is identifying all types of dataset grouping. I t's used for the personal optimal global optimal solution of the process, and observes all kinds of steps in the dataset.

II.LITERATURE SURVEY

According to Machine Learning (ML) is any process that results in accurate and reliable output of the process. The presented method is a detection in credit fraud and solves the process. The focus in the ML techniques is impossible in the real-time of the dataset collected [6]. The proposed method is Deep Learning (DL), which is a collect of the real-time processing and standard scalability of the process. However, the process is more time-consuming and expensive than the techniques.

Two-Stage Thresholding (TST) is technique is a simple technique he in the dataset of the classification process because use in credit fraud detection as an early method. The presented approach focuses on the two stages of classifying all kinds of datasets for easy prevention, and the technique is impossible in a more data-driven analysis process [7]. The proposed Method is a Synthetic Minority Over-sampling Technique (SMOTE) is particular time analyses is a more significant number of datasets and reducing the time complexity. However, the process a more occupied of space and sometimes a mismatching of the datasets.

Artificial Intelligence (AI) technology is used in fraud detection because the its early detection and easily prevention of the process. The presented techniques are a more quickly analysis of the dataset of the process and focus on the network connection of the data-driven. The process has a low range of the reliability of techniques and is impossible in real-time data analysis [8]. The proposed method DL is focused on the real-time data and is more secure of the network connection of the process. However, the process is a real-time data analysis in creating for impacting the environment of the techniques.

The ML techniques for analysis in the previous dataset were collected for the credit fraud detection process. The presented method is used in the individual dataset measure in the process, and a high level of performance is the network connection of the techniques, and the process is a more significant signal problem [9]. The proposed method is a Decision Tree that focuses on the network connection, which does not affect the environment and reduces the signal issue of the process. However, the processes are complex in building the network connection of the techniques.

The Naive Bayes method is a dataset classification because more occur in credit fraud in financial institutions. The presented techniques analyze the significant danger information, and the result is avoided. The network connection process is a slow level of performance and low range of efficiency [10]. The proposed method is an ensemble learning that is an autoencoder of the dataset for the network connection and process reliability. However, more inference is frequently a technique, and more power consumption.

The ML techniques are an analyze of the serious problem of the credit card detection process. The presented techniques collect the correct information, the process, and the classification. The method focuses on identifying the missed datasets and data-

driven process preservation, but the process is not standard for the stability of the techniques [11]. The proposed method is a variational automatic coding (VAE) for automatic detection in various datasets. However, the process is a higher-risk and costly technique.

The ML techniques are used a identify duplicate activities and enhance the imbalanced dataset of the process. The presented methods are based on AI technology and improve the overall performance in the network connection of the process. The overfitting occurs in the data-driven network connection of the process [12]. The proposed method Naive Bayes is a is more memory efficiency and high dimensional space, reduce for the overfitting datasets of the process. However, the process is a lack of interpretability and high computational of the techniques.

The Adaboost technique is a one of the ensembles learning method and based on the ML technology. The presented method is a weak classifier improve in the strong classifier of the process, and the dataset is more sensitive and high time-consuming range of the process [13]. The proposed method K- Nears Neighbour (KNN) is a process determine in the Neighbour node of the network connection of the process. However, the processes a high memory usage and slow prediction of the techniques.

The ML techniques for analysis in the previous dataset were collected for the credit fraud detection process. The presented method is used in the individual dataset measure in the process, and a high level of performance is the network connection of the techniques, and the process is a more significant signal problem [13]. The proposed method is Supervised Learning algorithm, which is a collect of the real-time processing and standard scalability of the process. However, the process is more time-consuming and expensive than the techniques.

The ML techniques are a detection of the credit fraud in early stages, and the prediction of the result is more efficiency. The presented method is a focus for the classification in the valid and invalid data of the process, but the ML techniques is a changing in the channel for the network connection of the process [14]. The proposed method is a Deep Neural Network (DNN) is a enhance the built in the network process and reduce in th e changing channel. However, the process is a more overfitting datasets and low scalability range of the process.

Two-Stage Thresholding (TST) is technique is a simple technique he in the dataset of the classification process because use in credit fraud detection as an early method. The presented approach focuses on the two stages of classifying all kinds of datasets for easy prevention, and the technique is impossible in a more data-driven analysis process [15]. The proposed method is Deep Learning (DL), which is a collect of the real-time processing and standard scalability of the process. However, the process is more time-consuming and expensive than the techniques.

The Naive Bayes method is a dataset classification because more occur in credit fraud in financial institutions. The presented techniques analyze the significant danger information, and the result is avoided. The network connection process is a slow level of performance and low range of efficiency [16]. The proposed method is Supervised Learning algorithm, which is a collect of the real-time processing and standard scalability of the process. However, the process is more time-consuming and expensive than the techniques.

Artificial Intelligence (AI) technology is used in fraud detection because the its early detection and easily prevention of the process. The presented techniques are a more quickly analysis of the dataset of the process and focus on the network connection of the data-driven. The process has a low range of the reliability of techniques and is impossible in real-time data analysis [17]. The proposed method DNN is a enhance the built in the network process and reduce in th e changing channel. However, the process is a more overfitting datasets and low scalability range of the process.

The ML techniques are an analyze of the serious problem of the credit card detection process. The presented techniques collect the correct information, the process, and the classification. The method focuses on identifying the missed datasets and data-driven process preservation, but the process is not standard for the stability of the techniques [18]. The proposed method is a Decision Tree making is the autoencoder of the dataset for the network connection and process reliability. However, more inference is frequently a technique, and more power consumption.

The proposed Method is a Synthetic Minority Over-sampling Technique (SMOTE), which is a particular time analysis for more significant datasets and reduces the time complexity. However, the process a more occupied of space and sometimes a mismatching of the datasets [19]. The proposed method, DL, is focused on real-time data and is more secure for the network connection of the process. However, the process is a real-time data analysis in creating for impacting the environment of the techniques.

The ML technique is one of the ensembles learning methods and is based on ML technology. The presented method is a weak classifier that improves the strong classifier of the process, and the dataset is a more sensitive and time-consuming process range [13]. The proposed method, K-Near Neighbour (KNN), is a process that determines the nearest Neighbour node of the network connection of the process. However, the processes have a high memory usage and slow prediction of the techniques.

III. PROPOSED METHODOLOGY

The proposed method is an SVM-DT, a narrow channel of the best solution in the process. Individual personal data is enhanced to a secure level, and a classification of the dataset is more efficient in stabilizing the process. SVM collected the dataset for classification of the binary type of data, and data select in the hyperplane for the margin of the dataset of the network connection of the process. The Decision Tree is a fraud detection process in measuring techniques and dataset classification. The presented method is a identify the misclassification of the data types, the probability of all kinds of nodes, and the gain of all the dataset classifications in the credit fraud detection process.

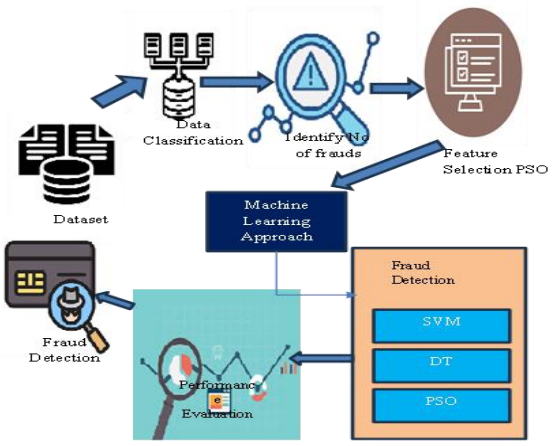

Fig. 1 Credit Card Fraud Detection Using SVM-DT

Figure 1 shows a dataset collected for the classification of binary data, with data selected in the hyperplane for the margin of the dataset of the n-process network connection fraud detection process uses measuring techniques and dataset classification. It is a narrow channel of the best solution in the process, and individual personal data is enhanced to a secure level, and a classification of the dataset is more efficient in the stability of the process. The process is the shortest path is identifying all types of dataset grouping. I t's used for the personal optimal global optimal solution of the process, and observes all kinds of steps in the dataset.

A. Dataset Description

The two days are a dataset transaction, with 492 frauds out of 284,807 transactions. Highly unbalanced dataset of the process; the positive class (frauds) accounts for 0.172% of all transactions. The first transaction in the dataset Feature 'Time' contains the seconds elapsed between each transaction. The transaction amount in the feature amount of the process; The feature of the method used in the cost-sensitive. Feature 'Class' is the response variable that takes the binary value is 0 or 1 in the process.

	A	B	C	D	E	F	G	H	I	J	K
1	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
2	0	-1.35981	-0.07278	2.536347	1.378155	-0.33832	0.462388	0.239599	0.098698	0.363787	0.090794
3	0	1.191857	0.266151	0.16648	0.448154	0.060018	-0.08236	-0.0788	0.085102	-0.25543	-0.16697
4	1	-1.35835	-1.34016	1.773209	0.37978	-0.5032	1.800499	0.791461	0.247676	-1.51465	0.207643
5	1	-0.96627	-0.18523	1.792993	-0.86329	-0.01031	1.247203	0.237609	0.377436	-1.38702	-0.05495
6	2	-1.15823	0.877737	1.548718	0.403034	-0.40719	0.095921	0.592941	-0.27053	0.817739	0.753074
7	2	-0.42597	0.960523	1.141109	-0.16825	0.420987	-0.02973	0.476201	0.260314	-0.56867	-0.37141
8	4	1.229658	0.141004	0.045371	1.202613	0.191881	0.272708	-0.00516	0.081213	0.46496	-0.09925
9	7	-0.64427	1.417964	1.07438	-0.4922	0.948934	0.428118	1.120631	-3.80786	0.615375	1.249376
10	7	-0.89429	0.286157	-0.11319	-0.27153	2.669599	3.721818	0.370145	0.851084	-0.39205	-0.41043
11	9	-0.33826	1.119593	1.044367	-0.22219	0.499361	-0.24676	0.651583	0.069539	-0.73673	-0.36685
12	10	1.449044	-1.17634	0.91386	-1.37567	-1.97138	-0.62915	-1.42324	0.048456	-1.72041	1.626659
13	10	0.384978	0.616109	-0.8743	-0.09402	2.924584	3.317027	0.470455	0.538247	-0.55889	0.309755
14	10	1.249999	-1.22164	0.38393	-1.2349	-1.48542	-0.75323	-0.6894	-0.22749	-2.09401	1.323729
15	11	1.069374	0.287722	0.828613	2.71252	-0.1784	0.337544	-0.09672	0.115982	-0.22108	0.46023
16	12	-2.79185	-0.32777	1.64175	1.767473	-0.13659	0.807596	-0.42291	-1.90711	0.755713	1.151087
17	12	-0.75242	0.345485	2.057323	-1.46864	-1.15839	-0.07785	-0.60858	0.003603	-0.43617	0.747731
18	12	1.103215	-0.0403	1.267332	1.289091	-0.736	0.288069	-0.58606	0.18938	0.782333	-0.26798
19	13	-0.43691	0.918966	0.924591	-0.72722	0.915679	-0.12787	0.707642	0.087962	-0.66527	-0.73798
20	14	-5.40126	-5.45015	1.186305	1.736239	3.049106	-1.76341	-1.55974	0.160842	1.23309	0.345173
21	15	1.492936	-1.02935	0.454795	-1.43803	-1.55543	-0.72096	-1.08066	-0.05313	-1.97868	1.638076
22	16	0.694885	-1.36182	1.029221	0.834159	-1.19121	1.309109	-0.87859	0.44529	-0.4462	0.568521

Fig. 2 Dataset of Feature Selection

B. Particle swarm optimization algorithm (PSO)

In the section offered PSO is a narrow channel of the best solution in the process, and individual personal data is enhanced a secure level, and more efficient in the stability of the process is a classification of the dataset. The process is the shortest path is identifying all types of the datasets grouping. I t's used for the personal optimal global optimal solution of the process, and observes all types of steps in the dataset.

Equation 1 eliminates misclassification, changing channels, observing individual data, and the quality interaction of environmental features.

$$S_{zm}(t+1) = s_{zm(t)} + y_1 R_1\left(Q_{zm}(t) - l_{zm}(t)\right) + y_2 R_2\left(p_{gd}(t) - y_{zm}(t)\right) \quad (1)$$

The equation 2 is a calculation in each process step; it's more effective in testing for false negatives in the dataset to detect the process.

$$Y_{zm}(t+1) = Y_{zm}(t) + S_{zm}(t+1) \quad (2)$$

Equation 3 is a feature selection for increasing the computation time and memory resources that are unavoidable for the classification techniques.

$$fd = \{p_{q1}, p_{q2}, p_{q3}, p_{q4}, \cdots, p_{qn}\} \quad (3)$$

Equation 4 eliminates the overhead and redundancy, and the data transmission process of the dataset improves in reliability of the process.

$$fd' = \{p_{q1}, p_{q2}, p_{q3}, p_{q4}, \cdots, p_{qn}\} \quad (4)$$

The PSO is the ability to adapt to new conditions based on adequate time and space computations. Let's assume the Q-number of the rough feature after elimination, and iteration, y-acceleration, z-particle, n-mass of the direction, Vzm-global optimal, and Yzm-personal optimal of the process.

C. Support Vector Machine-Decision Tree (SVM-DT)

The SVM-DT method is a collected the dataset for classification of the binary type of data, and data select in the hyperplane for the margin of the dataset of the network connection of the process. The present method verifies the wrong margin and hyperplane, identifying the misclassification for a process dataset. The process is based on ML techniques and is more accurate in performance. The method SVM-DT is a process of fraud detection in measuring techniques and dataset classification. The presented method is a identify the misclassification of the data types, the probability of all kinds of nodes, and the gain of all the dataset classifications in the credit fraud detection process.

Equation 5 various types of datasets classified in the data type of processing, with two separate types of variables in binary datasets.

$$p.n.r_{i(z.\emptyset(q_i)+b)\geq 1, i=1,2,\ldots,l} \quad (5)$$

Equation 6 is a classification error determined using a soft margin in the process techniques, and the wrong hyperplane is allowed in the dataset collected from the process because it stores the incorrect classification of the data.

$$p.n.r_{i(z.\emptyset(q_i)+b)\geq 1-\varepsilon i, =1,2,\ldots,l} \quad (6)$$

The above equation of the problem requires the use of sampling techniques, and the equation 7 which is a process where a dataset is classification correctly in the hyperplane of a surface in the dimensions of the process non-explicitly in the dataset methods.

$$0 \leq \beta_i \leq C, i = 1,2,\ldots,l \quad (7)$$

Equation 8 is a reduce in the financial loss, and the process measures the individual datapoint and enhances the data classification and hyperparameter in the process parameter.

$$p.n.r_i.(z.\emptyset(q_i) + b) \geq 1 - \varepsilon_{i,} i = 1,2,\ldots,l \ , \varepsilon_i \geq 0, i = 1,2,\ldots,l \quad (8)$$

The process is determined in the binary dataset, and the select of the hyperplane and hyperparameter of the process. Let's assume binary datasets, z-hyper plane, n-number of nodes, and q- q-quality of the classification of the process.

The algorithm is a find in the various types of nodes and decision making in the network of the nodes for a process.

Equation 9 shows the number of fraud nodes collected based on the process's fraud detection and the methods' dataset classification.

$$SLR = \frac{\sum_{j=1}^{x}(P_{sq})j}{\sum_{i=1}^{y}(P_{sq})i} \quad (9)$$

Equation 10 measures the process's impurity and the splitting of various nodes. It also reduces the total number of misclassifications in the process dataset.

$$P_q = \sum_{i=1}^{x}(P_{sq})i \quad (10)$$

Equation 11 reduces the cost of selecting the process's higher frequency values in the nodes and the distribution in the class nodes.

$$P_r = \left(\frac{Q}{Q+X}\right) * Q * P_{sq} \quad (11)$$

Equation 12 and 13 calculates the total number of misclassifications and labeled nodes and the distribution of class nodes in the process. The measure in the zero classification and minimize the misclassification of the dataset, and th probability of the process.

$$P_{q=Node} = P_s/\left(P_q + P_S\right) \quad (13)$$

$$P_{s=Node} = P_s/\left(P_q + P_S\right) \quad (12)$$

Equation 14 is the gain ratio of the various nodes in the child node and parent node of the decision-making process, and the more secure dataset in the

fraud section, due to the fault of the processes solved for the process.

$$P_z = \left(\sum_{i=1}^{n}(P_{ps})\, i\right)/n \qquad (14)$$

The process is a decision-making for the classification of the dataset, improving the eliminate of mismatching data and changing the channel of the process. Let's assume Cs-class probability, k-number of nodes, Psq-misclassification cost, and r-frequency class.

IV. RESULT AND DISCUSSION

This section evaluates the precision, recall, accuracy, time complexity, and FN score across various parameters and approaches. Furthermore, the proposed method can securely transfer personal data in Machine Learning using 2,84,808 data points in the attack dataset.

Table.1 Simulation Parameters

Simulation	Parameter Name
Dataset	Fraud detection Dataset
No of Dataset	2,84,808
Training Dataset	1,84,808
Testing Dataset	1,00,000
Language	Python
Tool	Juypter

As illustrated in Table 1, the simulation parameters were evaluated through 2,27,808 dataset nodes collected in the feature selection process. 1,27,208 is a training dataset for the process, and 1,00,00 is a testing dataset.

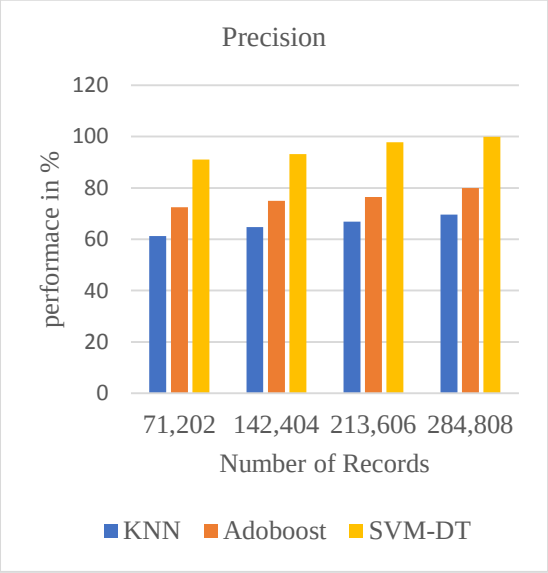

Fig. 3 Analysis of Precision Level

Figure 2 illustrates using precision-level analysis for secure health data exchange through Wireless Network technology. This review assesses previous methods, including KNN, Adaboost, Naïve bayes and contrasts them with the proposed SVM-DT method. The precision level of the performance ratings for these methods is 85.6,89.9, and 99.2, respectively, for the various performance levels in data protection.

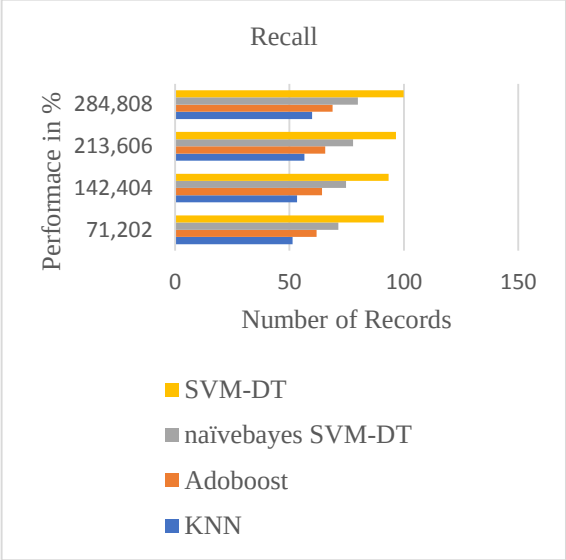

Fig. 3 Analysis of Recall

Figure 3 illustrates using recall analysis for secure health data exchange through Wireless Network technology. This review assesses previous methods, including KNN, TST, Adaboost and contrasts them with the proposed SVM-DT method. The recall of the performance ratings for these methods is 69.6,79.9, and 89.2, respectively, for the various performance levels data protection.

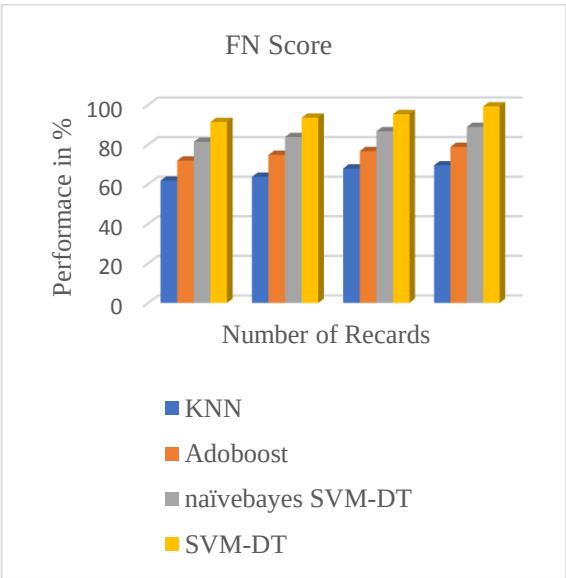

Fig. 4 Analysis of FN Score

Figure 4 illustrates using FN Score analysis for secure health data exchange through Wireless

Network technology. This review assesses previous methods, including KNN, Adaboost, Naïve bayes and contrasts them with the proposed SVM-DT method. The recall of the performance ratings for these methods is 65.6,79.9,88.5, and 99.2, respectively, for the various performance levels in data protection.

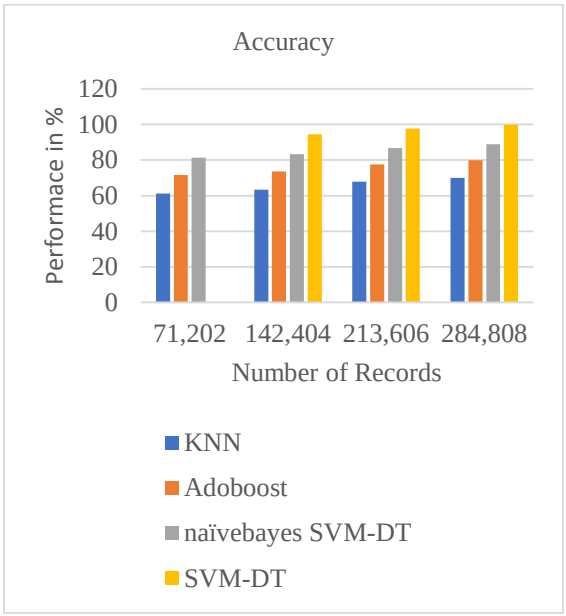

Fig. 5 Analysis of Accuracy

Figure 5 illustrates using Accuracy analysis for secure health data exchange through Wireless Network technology. This review assesses previous methods, including KNN, Adaboost, Naïve bayes and contrasts them with the proposed SVM-DT method. The accuracy of the performance ratings for these methods is 59.6,69.9,78.8 and 89.2, respectively, for the various performance levels in data protection.

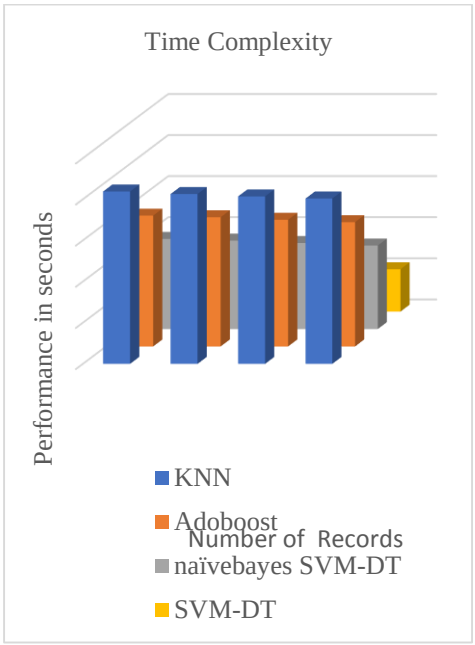

Fig. 6 Analysis of Time Complexity

Figure 6 illustrates using Time Complexity analysis for secure health data exchange through Wireless Network technology. This review assesses previous methods, including KNN, Adaboost, Naïve bayes and contrasts them with the proposed SVM-DT method. The time complexity of the performance ratings for these methods is 4.05,3.99, 2.78 and 1.32, respectively, for the various performance levels in data protection.

V. CONCLUSION

This study analysis the crucial role of Machine Learning techniques in enhancing threat detection, prevention, and response in credit card fraud detection operations. Security analysts respond to cyber threats quickly in many security incidents by minimizing false positive alerts. It improves performance measurement using F1 score, time complexity, recall, precision, and accuracy based on commonly used benchmark datasets to compare results. It enhances the capacity to identify fraud detection by compressing vast amounts of data into practical archives, using various learning models and DL-based detection algorithms for specific credit card fraud detection. Using frameworks like KNN, Adaboost, Naïve bayes and SVM-DT to compare long-term security data allows a quick and efficient response to critical security alerts. Improving accuracy by comparing performance estimates using the Fraud Detect 2013 dataset collected in the real world. In the future, to address the evolving fraud detection problem, we will focus on enhancing earlier threat predictions through a multiple deep learning approach to discovering the long-term patterns in history data, and the process has achieved a 92% accuracy in fraud detection.

REFERENCE

[1] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," in *IEEE Access*, vol. 10, pp. 39700-39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[2] Bin Sulaiman, Rejwan, Vitaly Schetinin, and Paul Sant. "Review of machine learning approach on credit card fraud detection." *Human-Centric Intelligent Systems* 2.1 (2022): 55-68.

[3] M. Adil, Z. Yinjun, M. M. Jamjoom and Z. Ullah, "OptDevNet: A Optimized Deep Event-Based Network Framework for Credit Card Fraud Detection," in *IEEE Access*, vol. 12, pp. 132421-132433, 2024, doi: 10.1109/ACCESS.2024.3458944.

[4] E. Ileberi, Y. Sun and Z. Wang, "Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost," in *IEEE Access*, vol. 9, pp. 165286-165294, 2021, doi: 10.1109/ACCESS.2021.3134330.

[5] I. D. Mienye and Y. Sun, "A Deep Learning Ensemble With Data Resampling for Credit Card Fraud Detection," in *IEEE Access*, vol. 11, pp. 30628-30638, 2023, doi: 10.1109/ACCESS.2023.3262020.

[6] I. D. Mienye and N. Jere, "Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions," in *IEEE Access*, vol. 12, pp. 96893-96910, 2024, doi: 10.1109/ACCESS.2024.3426955.

[7] I. Almubark, "Advanced Credit Card Fraud Detection: An Ensemble Learning Using Random Under Sampling and Two-Stage Thresholding," in *IEEE Access*, vol. 12, pp. 192079-192089, 2024, doi: 10.1109/ACCESS.2024.3519335.

[8] F. Khaled Alarfaj and S. Shahzadi, "Enhancing Fraud Detection in Banking With Deep Learning: Graph Neural Networks and Autoencoders for Real-Time Credit Card Fraud Prevention," in *IEEE Access*, vol. 13, pp. 20633-20646, 2025, doi: 10.1109/ACCESS.2024.3466288.

[9] Z. Xie and X. Huang, "A Credit Card Fraud Detection Method Based on Mahalanobis Distance Hybrid Sampling and Random Forest Algorithm," in *IEEE Access*, vol. 12, pp. 162788-162798, 2024, doi: 10.1109/ACCESS.2024.3421316.

[10] J. Jemai, A. Zarrad and A. Daud, "Identifying Fraudulent Credit Card Transactions Using Ensemble Learning," in *IEEE Access*, vol. 12, pp. 54893-54900, 2024, doi: 10.1109/ACCESS.2024.3380823.

[11] H. Tingfei, C. Guangquan and H. Kuihua, "Using Variational Auto Encoding in Credit Card Fraud Detection," in *IEEE Access*, vol. 8, pp. 149841-149853, 2020, doi: 10.1109/ACCESS.2020.3015600.

[12] H. Wang, W. Wang, Y. Liu and B. Alidaee, "Integrating Machine Learning Algorithms with Quantum Annealing Solvers for Online Fraud Detection," in *IEEE Access*, vol. 10, pp. 75908-75917, 2022, doi: 10.1109/ACCESS.2022.3190897.

[13] W. Ning, S. Chen, S. Lei and X. Liao, "AMWSPLAdaboost Credit Card Fraud Detection Method Based on Enhanced Base Classifier Diversity," in *IEEE Access*, vol. 11, pp. 66488-66496, 2023, doi: 10.1109/ACCESS.2023.3290957.

[14] H. Palivela *et al.*, "Optimization of Deep Learning-Based Model for Identification of Credit Card Frauds," in *IEEE Access*, vol. 12, pp. 125629-125642, 2024, doi: 10.1109/ACCESS.2024.3440637.

[15] R. San Miguel Carrasco and M. -Á. Sicilia-Urbán, "Evaluation of Deep Neural Networks for Reduction of Credit Card Fraud Alerts," in *IEEE Access*, vol. 8, pp. 186421-186432, 2020, doi: 10.1109/ACCESS.2020.3026222.

[16] S. K. Hashemi, S. L. Mirtaheri and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," in *IEEE Access*, vol. 11, pp. 3034-3043, 2023, doi: 10.1109/ACCESS.2022.3232287.

[17] M. Alamri and M. Ykhlef, "Hybrid Undersampling and Oversampling for Handling Imbalanced Credit Card Data," in *IEEE Access*, vol. 12, pp. 14050-14060, 2024, doi: 10.1109/ACCESS.2024.3357091.

[18] Alarfaj, Fawaz Khaled, et al. "Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms." *Ieee Access* 10 (2022): 39700-39715.

[19] Mienye, Ibomoiye Domor, and Yanxia Sun. "A deep learning ensemble with data resampling for credit card fraud detection." *Ieee Access* 11 (2023): 30628-30638.

[20] Varun Kumar, K. S., et al. "Credit card fraud detection using machine learning algorithms." *International journal of engineering research & technology (IJERT)* 9.7 (2020): 2020.