IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

# Financial Markets Fraud Detection and Prevention using Kernel Adatron Algorithm with Machine Learning

**Dr. Om Prakash Yadav[1], Mr. Ravi Kant Sahu[2], Mr. Anand Soni[3]**
om.26121@lpu.co.in[1], ravi.16920@lpu.co.in[2], anand.soni@bti.moe.bh[3]
Assistant Professor, School of Computer Science and Engineering[1,2]
Lovely professional University, Phagwara, Punjab, India
Sr. Lecturer, Commercial Studies Division Bahrain Training Institute[3]
Ministry of Education, Kingdom of Bahrain

**Abstract:** *The investors may have great interest to invest in stock market. Moreover, financial markets like stock markets are driven by explosive factors such as social media, micro blogs and news that make it hard to predict stock market index based on purely the historical data. The financial services industries that involve financial transactions are suffering from fraud- related losses and damages. Machine Learning (ML) is being rapidly adopted for a range of applications. It is important to begin considering the financial stability implications for every financial asset"s organization. Using the machine learning tools and techniques in the finance sector will become necessary because it will closely monitor nascent and rapidly evolving landscape, wherein data on usage are largely unavailable, and bereft of any analysis. Financial assets fraud has seriously affected investors' confidence in the stock market and economic stability. The huge economic losses incurred because of several serious financial fraud events and because of this the intelligent financial fraud detection has thus been the topic of recent advances. In recent years, several studies have used stock market and machine learning techniques to provide solutions to this problem. In this paper, we propose various a state of art fraud detection techniques such as classification, clustering, and regression. This study aims to identify the techniques and methods that give the best results that have been perfected so far. Stock markets can benefit if fraud identification and prevention can be incorporated by using machine learning algorithms.*

**Keywords:** Financial Market, Fraud Detection & Prevention, Machine Learning. Kernel Adatron.

## I. INTRODUCTION

**Financial Fraud** - The "Fraud", "corporate crime" and "white collar crime" are all the various terms used interchangeably when referring to economic crimes, where fraudulent activities has occurred. The financial services industry and the industries that contain financial transactions are burdened from fraud-related losses and damages. According to RBI,229 frauds on average happened every day in financial year, 2020-21, during which more than 83,000 banking frauds took place that syphoned off Rs 1.38 lakh crore. The recoveries during the same period amount to a little over Rs 1,000 crore, less than one per cent of the outflow. The RBI Annual Report 2021-22 pointed out that the private sector banks in India recorded the highest number of frauds in the financial year 2021-22. RBI said in its annual report. The total number of fraud cases in FY22 stood at 9,103. In comparison, the countstood at 7,359 last year. The total fraud amount stood at Rs 60,414 crore in FY22 while the amount was Rs 1,38,211 crore a year ago.[22]

The Office of Inspector General for the U.S. Department of Labor (DOL) estimates about $87 billion in fraudulent claims; some experts think losses could be in the hundreds of billions in 2021. Out-of-date online unemployment systems also heightened the most scandalous fraud of 2021 [3]. It takes 40 plus days to detect fraud for brick-and mortar financial institutions according to Javelin Strategy & Research. The Fraud also effects banks that provide

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

online payments service. For instance, 20% of customers change their banks after experiencing scams. So, the challenge for industry players is to implement real-time claim valuation and improve the accuracy of fraud detection[8].

Fraud in financial assets affects tremendously both the financial industry and everyday life. Fraud can reduce confidence in the industry, destabilize savings and affect the cost of living[9]. Financial institutions use a variety of fraud prevention models to addressthis problem. However, fraudsters are adaptive, and over time, they conceive several ways of intruding such protective models. Despite the best effort of financial institutions, law enforcement and government, financial fraud continues to grow. Fraudsters today can be a very inventive and intelligent fraternity[11]. This paper, seeks to carry out comparative analysis of financial fraud detection techniques, like machine-learning techniques, which playan important role in fraud detection, as it is often applied to extract and uncover the hidden truths behind very large quantities of data. Also, many modern techniques for detecting fraud are continually involving and they are applied to many areas due to the remarkable growth of fraudulent activities which effects on financial markets field in each year. Our objective is to point out their strength and weaknesses and also aim to identify the open issues of fraud analysis[4].

Financial asset fraud can also happen when corporations intentionally prepare financial statements that include misstated or misrepresented material to mislead stock market investors and regulators (1). According to Hajek and Henriques (2), the commontypes of financial statement fraud include omissions in financial records, falsification or manipulation of revenue, income, assets, expenses and other financial variables, and misrepresentation of management discussions and analysis. Financial asset fraud seriously affects both the investors and regulators, and it also causes huge losses in the economy andthe stock market and destroys the general public‟s confidence in the business environment [35]. In the past few years, several firms have been involved in financial statement fraud activity, which led to economic turmoil [13]. For example, Enron and other firms perpetrated financial fraud, which enormously affected the world economy and stock market [37] (3). According to Abbasi (4), in the ten largest bankruptcies in United States history, four companies were involved in major financial fraud. Beasley et al. (5) showed that firms that commit fraud, 28% were bankrupted in two years, and 47% were delisted from the stock exchange[14]. Therefore, financial statement fraud has attracted much concern from investorsand regulators.

### 1.1 Stock Market Fraud

- **Financial fraud - *Definition*** Fraud definition, according to the Association of CertifiedFraud Examiners (ACFE) “ACFE Association of Fraud Examiners Certificates”, fraud includes any intentional or deliberate act of depriving another of property or money by cunning, deception or other unfair acts [28].
- **Types of financial fraud** There are several types of financial fraud; we present here a brief description of some of the main types of fraud. *Insurance fraud* can occur at manypoints in the insurance process (e.g., application, eligibility, rating, billing, and claims), and can be committed by consumers, agents and brokers, insurance company employees, healthcare providers, and others [9, 24] *Securities and commodities fraud***,** the FBI [16] provides brief descriptions of some of the most prevalent securities and commodities frauds encountered today. According to another definition by CULS [10], “The Ponzi Scheme, The Pyramid Scheme, Market Manipulation, High Yield Investment Fraud, Prime Bank Scheme, Advance Fee Fraud, Hedge Fund Fraud, Commodities Fraud, Foreign Exchange Fraud, Broker Embezzlement and Late-Day Trading” are example of securities frauds include theft from manipulation of the market, theft from securities accounts, and wire fraud[19]. *Money Laundering* is the process by which criminals conceal or disguise the proceeds of their crimes or convert those proceeds into goods and services [45].

It allows criminals to inject their illegal money into the stream of commerce, thus corrupting financial institutions and the money supply and giving criminals‟ unwarranted economic power [27]. Gao and Ye [46] similarly define money laundering as the process by which criminals “wash dirty money” to disguise its illicit origin and make it appear legitimate and “clean.” *Financial statement fraud (corporate fraud),* financial statements are a company's basic documents to reflect its financial status [39]. It had an objective as.

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

- Fraud these statements to make the business more profitable
- Improvement of the performance of the actions
- Reduction of tax obligations

Attempt to exaggerate performance due to managerial pressure *Credit card fraud* is essentially of two types; application and behavioural fraud [15]. A fraudsters obtain new cards from issuing companies using false information or other people's information using the application fraud, where Behavioural fraud can be of four types: mail theft, stolen/lost card, counterfeit card and „card holder not present" fraud [33] *Mortgage Fraud* is a specific form of financial fraud that refers to the manipulation ofa property or mortgage documents. It is often committed to distort the value of a property for the purpose of influencing a lender to finance a loan for it [27]

## II. MACHINE LEARNING

In the recent years Machine Learning (ML) approach has received a lot of publicity to fraud detection and shifted industry interest from rule-based fraud detection systems to ML- based solutions. For detecting fraudulent activities in financial market"s a rule-based approach can be implemented which will ascertain for the on-surface and evident signals where the machine learning allows for creating algorithms that process large datasets with many variables and help find these hidden correlations between user behaviour and the likelihood of fraudulent actions[1].

The machine learning systems strength is compared to rule-based ones thus ensuring faster data processing and less manual work. For example, smart algorithms fit well with behaviour analytics for helping reduce the number of verification steps[23]. Prominent financial institutions, however, already use the ML technology to combat fraudsters. For example, MasterCard integrated machine learning and AI to track and process such variables as transaction size, location, time, device, and purchase data. The system assesses account behaviour in each operation and provides real-time judgment on whether a transaction is fraudulent[44].

The Fraud Scenarios and Detection

1. **Insurance Claims Analysis for Fraud Detection-** The most common issues are property damage, car insurance scams, and fake unemployment claims. For successful detection the basic requirement is a good dataset and carefully selected models are Fakeclaims and Duplicate claims and overstating repair cost[19].
2. **Medical Claims and Healthcare for Fraud Detection-** Healthcare and medical insurance is a rich area for fraud schemes due to the complexity and bureaucratic processes, which requires many approvals, verifications, and other paperwork are Up coding, abuse scams, Personal identity, Medical receipts and bills.
3. **Stock market fraud detection & Protection using Machine Learning -** The banks, regulators and investment firms are often involved in monitoring possible money laundering activity: They must detect and inform each other about suspicious activities. Anomaly detection is one of the common antifraud approaches in data science. It is based on classifying all objects in the available data into two groups: normal distribution and outliers. By analysing these parameters, anomaly detection algorithms can answer the following questions:
    - Do clients access services in an estimated way?
    - Are user actions normal?
    - Are transactions typical?
    - Are there any inconsistencies in the information provided by users?

There are two types of machine learning approaches that are commonly used in anti- fraud systems: *unsupervised and supervised machine learning[20]*.

Supervised learning entails training an algorithm using labelled historical data[24]. In this case, the goal of training is to make the system predict these variables and the existing datasets which already have target variables marked in future data. Unsupervised learning models process unlabelled data and classify it into different clusters detecting hidden relations between variables in data items. Both the supervised and unsupervised styles combine to build robust fraud detection systems by:

1. Collecting data
2. Labelling data

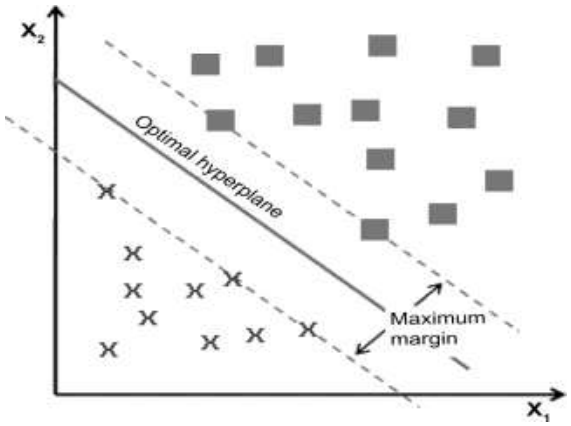
IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

3. Prediction accuracy
4. Training a supervised model.
5. Ensemble models.
6. Setting an express verification.

## III. SUPERVISED FRAUD DETECTION METHODS

In 1995 Vladimir Vapnik has first introduced SVM algorithm, it has been one of the most popular methods for classification because: it‟s a simple model, the use of kernel functions and the convexity of the function to optimize [10]. A support vector machine(SVM) is a supervised machine learning model that uses a non-probabilistic binary linear classifier to group records in a dataset[34].

The main idea behind the SVM is to construct a hyperplane as a decision dimension which maximizes the margin of separation between the positive and negative examples in a data set [38]. This induction principle is based on the fact that the error coefficient of the test data, that is, the coefficient of the generalization error, is limited by the sum of the coefficientof the training error, and this term depends on the dimension [42] .The performance of a support vector machine (SVM) depends highly on the selection of the kernel function type and relevant parameters [32]. SVM classifiers have been used for image de- noising, multi-class sentiment classification, or even for online suicide prevention.




We adopt binary SVM for classification [36] of manipulated samples where $Y \in \{-1,+1\}$ (i.e. 1 represents a manipulated sample). The main idea behind SVM is finding the*hyperplane* that maximizes the marginaldistance (i.e. sum of shortest distances) to data points in a class[25]. The

samples in input space are mapped to a feature space using a kernel function to find the *hyperplane*. We use the linear kernel in our experiments (other widely used kernels for SVMsare polynomial, radical basis function (RBF) and sigmoid [42].

SVMs with the main idea is of minimizing a regularized risk function R and maximizing the margin of separation between classes (Fig. 1) by solving Equation

$$w^* = arg \min_{w \in R^D} F(w) := \frac{1}{2} \left||w|\right|^2 + CR(w) \qquad (1)$$

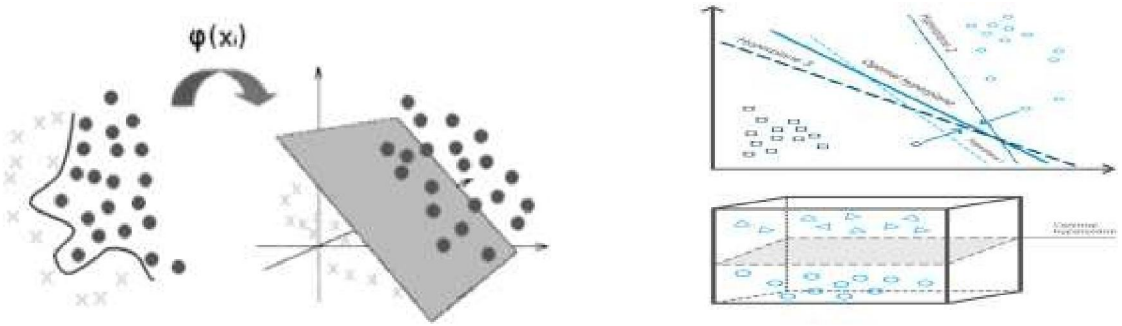
where w is a normal vector to the separating hyperplane, $1/2 \|w\|^2$ is a quadratic regularization term and C > 0 is the fixed constant that scales the risk function and equation 1 is called the primal formulation.[21] By using Lagrange multipliers, the primal formulation can be presented in its dual form:

$$L(\alpha) = \arg max \left\{ \sum_{i=0}^{n} \alpha_i - \sum_{i=0}^{n} \sum_{j=0}^{n} \propto_i \propto_j Y_i Y_j K(X_i, X_j) \right\} \qquad (2)$$

$$\text{subject to } 0 \leq \propto_i \leq C \text{ and } \sum_{i=1}^{n} \propto_i Y_i = 0$$

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

where $C$ is a fixed constant,$(X_i, Y_i)^n_{i=1}$ is a training set, ai are Lagrange multipliers, K(X , X ) is the value of the kernel matrix defined by the inner product $\langle X_i, X_j \rangle$ (when a linear kernel K is used) and Y ∈ {±1} is a class label[5]. The dual formulation has the same optimal values as the primal, but the main advantage of this representation is the use of the "kernel trick"




Since SVMs can only classify data in a linear, separable feature space, the role of the kernel function is to induce such feature space by implicitly mapping the training data into a higher dimensional space where data is linearly separable[40].

### 3.1 The Kernel Adatron Algorithm

The Adaptive Perceptron algorithm (or Adatron) was first introduced by J. K. Anlauf and M. Biehl in 1989 for training linear classifiers[5]. This algorithm was proposed as a method for calculating the largest margin classifier. The Adatron is used for on line learning perceptrons and guarantees convergence to an optimal solution, when this exists [29]. To implement KA algorithm, it is necessary to calculate the dot product w·X , where X is the set of training points and w denotes the normal vector to the hyperplane that divides the classes with a maximum margin (Fig. 1). Since the kernel K is related to the high-dimensional mapping φ(X ) by equation

$$K(X_i, X_j) = \varphi(X_i) \cdot \varphi(X_j) \quad (3)$$

where the normal vector w to the separating hyperplane, can be expressed as

$$W = \sum_{i=1}^{n} \alpha_i Y_i \varphi(X_i), \quad (4)$$

then, by using the lineal kernel K, the dot product can be expressed as

$$z_i = \sum_{j=1}^{n} \alpha_j Y_j K(X_i . Y_j), \quad (5)$$

To update the multipliers, a change in α must be proposed to be evaluated. The change can be calculated as follows

$$\delta\alpha = \eta(1 - \gamma_i), \quad (6)$$

$$\gamma = Y_i z_i, \quad (7)$$

where η is the step size and $\delta\alpha_i$ is the proposed change to $\alpha_i$. If $\alpha_i + \delta\alpha_i \leq 0$ it would result in a negative $\alpha_i$. To avoid this problem, $\alpha_i$ is set to 0. Otherwise, update $\alpha_i \leftarrow \alpha_i + \delta\alpha_i$ . The bias b can be obtained as follows:

$$b = \frac{1}{2}(\min(z_i^+) + \max(z_i^-)), \quad (8)$$

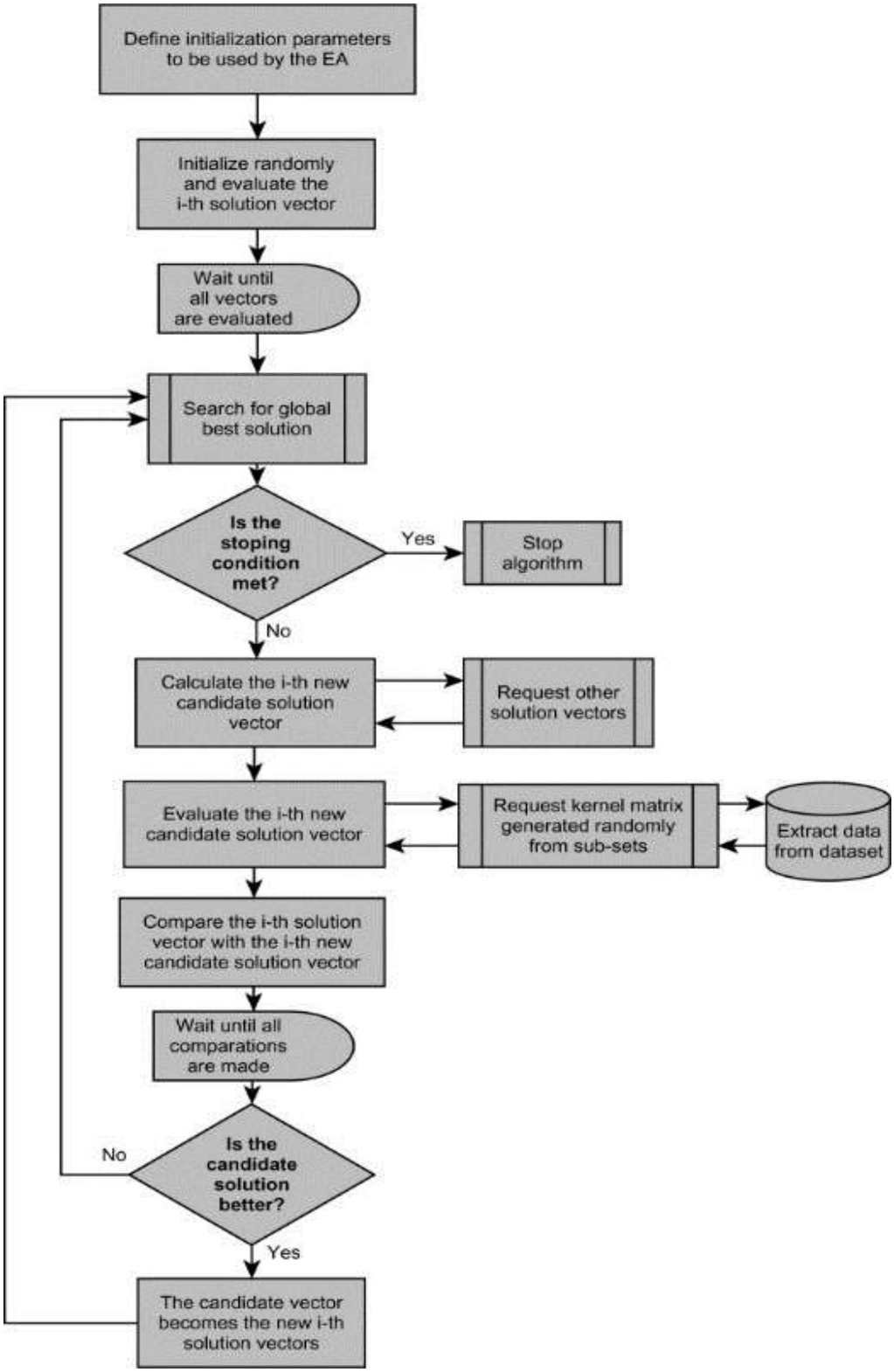
Where $Z_i^+$ is the patterns with class label +1 and those with class label −1.

The pseudocode is described briefly in Algorithm 1

1. Initialize α = 1.
2. **repeat**
3. For (X , Y ) calculate z with Equ - 5
4. Calculate γ with Equ - 7
5. Calculate δα with Equ - 6
6. **if** (α + δα ) ≤ 0 **then**
7. α = 0
8. **end if**

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

9. **if** $(\alpha + \delta\alpha) > 0$ **then**
10. $\alpha = \alpha + \delta\alpha$
11. **end if**
12. calculate b with Equ - 8
13. **until** The stopping criteria is met.



IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

Evolutionary algorithms (EA) or computing is a subfield of artificial intelligence that includes a range of problem-solving techniques based on principles of stack values evolution. The KA algorithm requires the α value to be adjusted through iterations. In this approach, the adjustment is made using EA.

This type of algorithm was chosen as an optimization method because they are easy to implement, to parallelize and have shown good results in diverse areas such as computer vision, image processing and path planning [3/36].

The basic idea behind the proposed algorithms is to use a "divide and conquer" strategy, where each individual in the population of the EA. On each variant of the proposed algorithm, individual x (particle, vector or bee) represents a D dimensional vector composed of multipliers to be optimized over iterations by the EA.

The fitness function f() to be used by the EA is described by Equation 9:

$$f(x) = \text{abs}(1 - \Theta) \quad (9)$$

where Θ is the margin between classes of the hyperplane, which can be estimated as follows:

$$b = \frac{1}{2}(\min(z_i^+) + \max(z_i^-)), \quad (10)$$

The value z can be obtained with equation 5. The values of z can be divided into and depending on their class label, $z_i^+$ and $z_i^-$, respectively. The KA algorithm has the implementation ease of the Adatron model and can find a solution very rapidly compared to conventional methods like kernel-perceptron and SVM [17]. The algorithm comes with all the theoretical assurances provided by the support vector theory for the large margin classifiers, and the convergence properties studied in the statistical learning literature [18]. However, the algorithm uses basic operations and has a complexity of O (n). Because of this, the algorithm has been modified so it can be trained using an EA with a computationally more attractive fitness function.

## IV. SUPERVISED FRAUD PREVENTION METHODS

Prevention of fraud is another area in which machine learning is evolving both as workflows and outcomes, thereby facilitating organizations to stay ahead of technologically sophisticated criminals. Meanwhile, fraudsters‟ techniques are growing increasingly technologically advanced. The entire software and hardware needed for fraud commission at scale are on sale, often for disturbingly low prices[31]. Contemporary online businesses are facing an increasingly sophisticated enemy that attacks, responds, and changes tactics extremely quickly. With machine learning, companies can stay ahead. The processes of preventing fraudsters are costly and time-consuming, but fighting fraudsters is important to ensure a well-functioning society. Good corporate governance can make sure that the firm´s top management don‟t misappropriate assets and manage results in an unethical manner, while internal processes can help preventing employees stealing from the firm [26]. The use of artificial intelligence and machine learning is now a worldwide phenomenon and companies in developing territories are investing in these technologies more compared to companies in developed territories [30].

Machine learning is also a commonly used method for fraud detection and prevention.Machine learning algorithms discovers patterns in big data and the process is much more efficient compared to humans doing the same task. With the information acquired throughout the process, it is easier to predict and prevent frauds. Machine learning can be divided into supervised and un-supervised machine learning. Supervised machine learning is the more commonly used by these two[43]. What differs these two machine learning methods, is thatin supervised machine learning guidelines and what conclusions it should come up with are given to the algorithms, this requires that possible outputs are already known. Unsupervised machine learning is instead identifying complex processes and patterns without any guidelines and human intervention, which can help with solving problems that humans normally couldn‟t do [12]

Another method is meta-learning. Meta-learning is a form of machine learning which uses information acquired through data mining or machine-learning with the purpose to increase the quality of results obtained in future applications, also called learning-to-learn. It differs from machine-learning, since meta-learning provides a way to learn about the process itself and by that, also providing knowledge about which features and algorithms that can be most efficiently applied [2]

The reasons why machine learning is an appropriate technology for fraud prevention are:

1. Fraud hides within an immense amount of data

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

2. Fraud happens fast
3. Fraud is unique
4. Fraud constantly changes
5. Fraud looks "clean"

The benefits of machine learning for fraud prevention-

1. Highly accurate results
2. Reduce the need for manual review
3. Fewer false positives, through sophisticated behavioural analysis
4. Fraud prevention ability without adversely affecting the user experience
5. Lower operational costs than other approaches
6. Frees up teams" time to focus on more strategic tasks
7. Can be automated
8. Adapts quickly

## V. RESULT ANALYSIS

This work presents a series of parallelized algorithms based on the KA algorithm as fitness function combined with Artificial Bee Colony (ABC), micro-Artificial Bee Colony (μABC), Differential Evolution (DE) and Particle Swarm Optimization (PSO), in order to solve the SVM learning problem. The data to classify was taken from the Interdisciplinary Computing and Complex stock market data Prediction Benchmarks Repository and five other datasets from diverse fields that are commonly used to test large-scale classifiers; the datasetsare briefly described in Tables 1 and 2.

**Table 1**

| DATASET | DIMENSION | DENSITY |
|---|---|---|
| TATA Motors | 99757 | 0.08% |
| Maruti Suzuki | 20707 | 0.23% |
| Hindustan Motors Limited | 47236 | 0.16% |
| Ashok Leyland | 20958 | 0.23% |
| Mahindra & Mahindra | 804 | 25.00% |

**Table 2**

| UNIFORM | Ω | DIMENSION | DENSITY |
|---|---|---|---|
| Length | 7 | 300 | 86.04 % |
| | 8 | 340 | 86.98 % |
| | 9 | 380 | 88.79 % |
| Frequency | 7 | 300 | 87.24 % |
| | 8 | 340 | 87.07 % |
| | 9 | 380 | 89.17 % |

From the PSP dataset, only the subsets discretized with uniform length and uniform frequency, with vehicle sizes ranging from 7 to 9, were used for training and generalization because of their density and dimensionality. The values used to train the SVM with each EA were obtained by running PSO on each variant of the algorithm to determinate the optimal values. This is not to be confused with the PSO variant that uses KA to classify data. The following values were used by the EAs while using the large-scale datasets:

- The μABC version used: RF = 0.0001, C = 0.0001, FCR = 0.0001 and maximum of 5 attempts before abandoning a food source.
- The ABC version used: C = 2, ϕ values ranging between [−2, 2], 5 food sources and a maximum of 9 attempts before abandoning a food source.
- The DE algorithm used: C = 2.38958, F = 1.87016 and CROV = 0.9 and 6 vectors.
- The PSO algorithm used: v = 1.49684, w = 1.18472, w = 0.000511895, c = 1.03971 c = 1.48063, C = 6.74659 and 15 particles.

For the PSP dataset, the following values were, used by the EAs:

- The μABC version used: RF = 0.001, C = 0.0001, FCR = 0.001, with maximum of 5 attempts before abandoning a food source and a maximum of 25 iterations as stopping condition.
- The ABC version used: C = 5, ϕ values ranging between [−2, 2], 8 food sources, a maximum of 9 attempts before abandoning a food source and a maximum of 20 iterations as stopping condition.
- The DE algorithm used: C = 2.65435, F = 0.719909 and CROV = 0.1, with 6 vectors and a maximum of 23 iterations as stopping condition.

IJARSCT
ISSN (Online) 2581-9429
nternational Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

- The PSO algorithm used: v = 0.1, w = 0.0494229, w = 0.0001, c = 1.13755 c = 0.11384, C = 3.5 with 10 particles and a maximum of 30 iterations as stopping condition.

The C value in SVM has two main purposes: it functions as constant that scales the risk function for the primal formulation in Equation 1 and it limits the values that any α can take in the dual formulation in Equation 2. As stated in Section the Kernel Adatron Algorithm, the KA algorithm has appealing advantages such as the simplicity of implementation of Adatron and the capability of working in high-dimension feature spaces to construct a large margin hyperplane. But the main concern of implementing the original KA approach is working with the kernel matrix, since its computational complexity is of $O(d*n^2)$, where d is the maximum number of non-zero features in any data vector of the training subset and n is the number of training samples. if it is treated as a divide and conquerproblem the computational complexity is reduced, at worst case scenario, to $O(d*n^2_{ts}/t)$, where t is the number of threads

## VI. CONCLUSION AND FUTURE OUTLOOK

It is important to consider that one of the peculiarities of internal fraud is that the person, or persons, that commit fraud often sees it as a victimless crime and can neither visualize any person who will be directly harmed [30]. This would also explain why the main perpetrators of economic crimes are internal actors, including human resources fraud (81%), insider trading (75%), asset misappropriation (75%), accounting fraud (74%) and also procurement fraud (73%) [30]. Current development of innovative and modern fraudand risk detection methods includes machine learning algorithms, data mining and meta- learning. They are all useful means accessible in risk and fraud prevention and detection systems [31; 30]. With that said, these methods will not be utilized in an efficient manner, if the organizational culture emphasize dishonest actions from perceived pressures, opportunities or rationalization [7].

Furthermore, when connecting technological advancement with organizationalculture, we find that future research should investigate this connection in more detail. For example, the aspects of the principal-agency theory could in future be applied to answer the questions regarding the possible situations of asymmetric information and conflicts of interest between management and stakeholders. Moreover, the practical business implications given from our findings, is that a combination of further development in organizational culture and more advanced technologies are the key to reduce the risk of fraud.

## REFERENCES

**[1].** Abbasi, Ahmed, Conan Albrecht, Anthony Vance, and James Hansen. 2012. "Metafraud: A Meta-Learning Framework for Detecting Financial Fraud." MIS Quarterly 36, no.4: 1293- 1327.

**[2].** Abbasi, Albrecht, Vance and Hansen (2012). Metafraud: A meta-learning framework for detecting financial fraud. MIS Quarterly, 36(4), 1293-1327.

**[3].** https://www.fraud-magazine.com/article.aspx?id=%204295016799.

**[4].** Analysis, F., Wireshark, U., In, T., & Internet, I. (2019). International Journal of Arts and Science Research frame analysis using wireshark and TOPAS in industrial internet of thinks. 6(1), 4–11.

**[5].** Anlauf JK, Biehl M. The Adatron: An adaptive perceptron algorithm. Europhysics Letters. 1989;10:687.

**[6].** Beasley, Mark S., Joseph V. Carcello, Dana R. Hermanson, and Terry L. Neal. 2010. "Fraudulent Financial Reporting: 2998-3007; An Analysis of U. S. Public Companies Research." New York: Committee of Sponsoring Organizations of the Treadway Commission (COSO). https://www.coso.org/Documents/COSO-Fraud-Study-2010-001.pdf

**[7].** Boston, W. (2015). Volkswagen Emissions Investigation Zeroes on Two Engineers. the Wall Street Journal 5.10.2015. Available: https://www.wsj.com/articles/vw-emissions-probe-zeroes-in-on-two-engineers-1444011602.Retrieved:3.3.2019.

**[8].** Campbell, C. (n.d.). The Kernel-Adatron Algorithm : a Fast and Simple Learning Procedure for Support Vector Machines. Systems Engineering.

**[9].** Coalition against Insurance Fraud, "Learn about fraud", http://www.insurancefraud.org/learn_about_fraud .htm.

**[10].** Cortes C, Vapnik V. Support vector networks. Machine Learning. 1995;20:273–97.
**[11].** CULS, Cornell University Law School, White-Collar Crime: an overview, http:// topics.law.cornell.edu/wex/White-collar_crime(2009).
**[12].** Data Science (2017). Supervised vs. Unsupervised Machine Learning. Available: https://www.datascience.com/blog/supervised-and-unsupervised-machine-learning- algorithms. Retrieved: 08.03.2019.
**[13].** De Castro, P. A. L., & Teodoro, A. R. B. (2019). A Method to identify anomalies in stock market trading based on Probabilistic Machine Learning. Journal of Autonomous Intelligence, 2(2), 42. https://doi.org/10.32629/jai.v2i2.44.
**[14].** Dong, Wei, Stephen Shaoyi Liao, Bing Fang, Xian Cheng, Zhu Chen, and Wenjie Fan. 2014. "The Detection of Fraudulent Financial Statements: An Integrated Language Model." Proceedings of 19th Pacific Asia Conference on Information Systems, June 24-28, 2014, Chengdu, China. Atlanta, GA: Association for Information Systems AIS eLibrary, 383. https://aisel.aisnet.org/pacis2014/383.
**[15].** Donning, H., Eriksson, M., Martikainen, M., & Lehner, O. M. (2019). Prevention and Detection for Risk and Fraud in the Digital Age-the Current Situation. ACRN Oxford Journal of Finance and Risk Perspectives, 8, 86–97.
**[16].** FBI, Federal Bureau of Investigation, Financial Crimes Report to the Public Fiscal Year, Department of Justice, United States, http: //www.fbi.gov/publications/ financial/fcs_report2007/financial_crime _2007.htm(2007) .
**[17].** Frieß TT, Cristianini N, Campbell C. The Kernel-Adatron algorithm: A fast and simple learning procedure for support vector machines; Proceedings of the Fifteenth International Conference on Machine Learning; Morgan Kaufmann; 1998.[33]
**[18].** Frieß TT, Harrison R. The Kernel-Adatron with bias unit: Analysis of the algorithm. 1998.
**[19].** Golmohammadi, K., & Zaiane, O. R. (2012). Data mining applications for fraud detection in securities market. Proceedings - 2012 European Intelligence and Security Informatics Conference, EISIC 2012, 107–114. https://doi.org/10.1109/EISIC.2012.5
**[20].** Hajek, Petr and Roberto Henriques. 2017. "Mining Corporate Annual Reports for Intelligent Detection of Financial Statement Fraud: A Comparative Study of Machine Learning Methods." Knowledge-Based Systems 128: 139-52.
**[21].** Haykin S. Neural networks: A comprehensive foundation. 3ed. Prentice-Hall; 2007.
**[22].** https://m.rbi.org.in/scripts/AnnualReportPublications.aspx?Id=1348
**[23].** International, A. (2017). A Design and Development as A system for Machine-to- Machine ( M2M ) working under vision motivated by the Internet of Things ( IOT ). XIII.
**[24].** J.L. Kaminski, Insurance Fraud, OLR Research Report, http: //www.cga.ct.gov/2005/ rpt/2005-R-0025.htm. 2004.
**[25].** Menon AK. Large-scale support vector machines: Algorithms and theory. Research Exam, University of California; San Diego: 2009. pp. 1–17.
**[26].** Murphy, P, R. and Free, C. (2016). Broadening the Fraud Triangle: Instrumental Climate and Fraud. Behaviroural research in accounting. 28(1), 41-56.
**[27].** Ngai E, Hu Y, Wong Y, Chen Y, and Sun X The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature Decision Support Systems Volume 50, Issue 3, p559-569 (2011).
**[28].** Online Available: http://www.acfe.com/uploadedfiles/acfewebsite/content/ documents/ rttn- 2010.pdf.
**[29].** Opper M. Learning times of neural networks: exact solution for a perceptron algorithm. Physical Review A. 1988;38:3824.
**[30].** PWC (2018). Pulling fraud out of the Shadows.Available: https://www.pwc.com/gx/en/forensics/global-economic-crime-and-fraud-survey-2018.pdf. Retrieved: 20.02.2019.
**[31].** Rezaee, Zabihollah. 2005. "Causes, Consequences, and Deterrence of Financial Statement Fraud." Critical Perspectives on Accounting 16: 277-98.
**[32].** S. Yin , J. Yin , Tuning kernel parameters for SVM based on expected square distance ratio, Inf. Sci.

(Ny). 370–371 (2016) 92–102.

**[33].** S.Bhattacharyya, S.Jha, K.Tharakunnel, J.C. Westland, Data mining for credit card fraud: A comparative study, Elsevier, Decision Support Systems, Volume 50, Issue 3, p602-613(2011).

**[34].** Sadgali, I., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. Procedia Computer Science, 148(Icds 2018), 45–54. https://doi.org/10.1016/j.procs.2019.01.007.

**[35].** Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. International Journal of Financial Studies, 7(2). https://doi.org/10.3390/ijfs7020026.

**[36].** Shi Y, Eberhart RC. Empirical study of particle swarm optimization; Proceedings of the 1999 Congress on Evolutionary Computation; pp. 1950–99.

**[37].** Tang, X. B., Liu, G. C., Yang, J., & Wei, W. (2018). Knowledge-based financial statement fraud detection system: Based on an ontology and a decision treet. Knowledge Organization, 45(3), 205–219. https://doi.org/10.5771/0943-7444-2018-3-205.

**[38].** V. Vapnik, The nature of statistical learning theory, 1995.

**[39].** W.H. Beaver, Financial ratios as predictors of failure, Journal of Accounting Research 4 p71– 111. (1966).

**[40].** West and Bhattacharya (2015). Intelligent financial fraud detection: A comprehensive review.

**[41].** Computers & Security, 57, 47-66.

**[42].** Y. Jin, R.M. Rejesus, B.B. Little, Binary choice models for rare events data: a crop insurance fraud application, Applied Economics Volume 37, Issue 7, p841–848. (2005).

**[43].** Y. Xu , L. Wang , P. Zhong , A rough margin-based ν-twin support vector machine, Neural Comput. Appl. 21 (2011) 1–11.

**[44].** Yadav, O. P. (2018). Intelligent Grid to Autonomous Cars and Vehicular Clouds in Internet of Vehicles. 7(2), 7–13.

**[45].** Yadav, O. P. (2019). Study on the Performance and Efficiency of Multi-Organize Filtering Framework with A Large Arrangement of SIP Security Based VoIP Domain. 3085(05), 1101– 1106.

**[46].** Yadav, O. P., & Singh, R. P. (2018). INTERNET OF THINGS ( IOT ) SECURITY ISSUE IN WIRELESS SENSOR NETWORK ( WSN ) WITH RADIO FREQUENCY IDENTIFICATION ( RFID ). 1–7.

**[47].** Z.Gao, M.Ye, A framework for data mining-based anti-money laundering research, Journal of Money Laundering Control 10 (2), p170–179.

## AUTHOR'S DETAILS

- Dr. Om Prakash Yadav, Assistant Professor in School of Computer Science and Engineering, Lovely Professor University. He worked ISL Engineering College, Sri Sarada Institute of Science and Technology and Grahmbell P.G. College, Hyderabad, Telangana. He has 16 years of teaching. He has published 16 International and 4 National Journals. He has participated and papers presented in in various National and International conferences. He had received Ph.D. degree from SSSTUMS, Bhopal, M.Tech from IETE, New Delhi and M.C.A from SMU.

- Mr. Ravi Kant Sahu, Assistant Professor, in School of Computer Science and Engineering, Lovely professor University. He has 10 years of teaching experience. He is awarded with M.Tech(CSE-Mobile Computing) from NIT Hamirpur (H.P.) [Gold Medalist] and B.Tech (I.T.) from Uttar Pradesh Technical University degrees. He is Oracle Certified Associate Java Programmer and qualified UGC NET and GATE exams. He has published 15+ International Journals. and participated in 5 National/International Level/ IEEE conferences.