Computational Intelligence for Financial Fraud Detection under Internet of Things Environment: Techniques, Opportunities and Challenges

Hossam Eldin M. Abd Elhamid[1], Wael Khalifa[2],
Mohamed Roushdy[3], Abdel-Badeeh M. Salem[4]

Abstract – **In this paper we introduce the financial problem of the fraud detection in credit cards and how it can be detected using computational intelligent techniques and Internet of Things (IOT). Moreover, we will present recent results and challenges.**

Keywords – **Machine Learning, Internet of Things IOT, Credit cards Fraud Detection.**

I. INTRODUCTION

With the increase of portable trade and the more usage of IoT, the more financial fraud could occur. The most common way of financial fraud could happen via credit card; and it could be with the presence of the credit card or without it. There are several techniques to reduce and capture this fraud, by using searching classification and implementing clustering most recent Known Techniques. Analyzing the points of interest and the restrictions, a model structure and full procedure try dependent on real budgetary exchange dataset then the creation of results and correlation with customary AI and profound learning.

II. RELATED WORKS

V. Sharma et al. suggested an approach that depends on the support of trust and safeguarding of protection rules in the social IoT framework; to forestall extortion in IoT a model was introduced that depends on intellectual signs which gives a Smart Detecting Model to Oddities.

Van Wyk Hartman, technique works by recognizing the misrepresentation in the distribution organize, the framework plans development and field examination to fix it [1].

Lasselare et al. introduced a Visa misrepresentation discovery framework that utilizes a few fundamental highlights to play out the recognition system [2] that depends on the purchasing conduct of the client. The significant boundaries used are regime, recurrence, and financial levels. These properties are utilized to anticipate corruption. However, the creator was restricted distinctly on Visa extortion type and not a worldwide framework. A standard-based extortion identification technique that gave colossal enhancements in the discovery procedure. This proposes to be an ongoing framework that has accomplished in a Turkish Bank. A cost-touchy charge card misrepresentation identification technique that utilizes the Base Hazard classifier [4]. The producer claims to give sensible perspectives on the monetary profits and losses happening because of fraud detection. A technique that focuses on giving effective misleading identification utilizing imbalanced information [5], this arrangement is viewed as an amazingly deficient and imbalanced information condition for playing out the extortion location process.

It is proved from the results that the recent solutions that can deal with the fraud systems, doesn't involve the technologies of Big Data as should be, none of them include the Internet of Things (IoT) or even try to discuss their way that they are using in collecting data. There are only few articles that used MapReduce and Hadoop but rarely some of them involved the concept of IoT.

.

III. INTERNET OF THINGS (I OT)

The unauthorized use of mobile transaction is called monetary corruption under IoT condition. Such unauthorized use can happen through data fraud or card taking to get cash falsely, because of that, it is viewed as the quickly developing issue through the rise of cell phone and online change administrations. In the real world, it is needed to highly take precautions since financial fraud causes financial loss.

Money related extortion under the IoT condition is the quickly developing issue since the mobile channel can facilitate any type of payment. As a result, the increasing in mobile commerce and the expansion of the IoT environment make the financial fraud in mobile payment more common. Above 87 percent of exporters bolster either a portable webpage or a versatile application for web-based shopping or both of them [6]. Also, supporting the mobile wallets builds the general utilization of money related to misrepresentation under the IoT condition. As a result, versatile installments under the IoT stage have reached $194.1 billion out of 2017, and portable vicinity installments additionally rose to $30.2 billion of every 2017, contrasted with $18.7 billion out of 2016[7].

Consequently, financial fraud, which typically happens under the IoT environment, is the most regular sort that includes assuming or changing acknowledgment card data. Furthermore, to recognize the issue of quickly producing extortion under the IoT condition, money related foundations utilize different misrepresentation counteraction devices like constant credit approval, address confirmation system, card check worth, the positive and the negative records, etc. [8]. In any case, existing detection systems rely upon characterized standards or scholarly records, which makes it hard to find new

[1] Hossam Eldin M. Abd Elhamid, Wael Khalifa & Abdel-Badeeh M. Salem is with Computer Science Dept. Faculty of Computers & Information Sciences, Ain Shams University, Cairo, Egypt E-Mail: hossam.abdelhamid@cis.asu.edu.eg, wael.khalifa@cis.asu.edu, absalem@cis.asu.edu.eg

[2] Mohamed Roushdy Is With Future University in Egypt, New Cairo, Egypt. E-Mail: Mohamed.Roushdy@fue.edu.eg

assault designs. In this manner, different techniques for utilizing AI and counterfeit neural systems have endeavored to catch new money related misrepresentation.

Machine learning is a field that machines learn ideas utilizing information, using statistical analysis to anticipate and arrange and input information as an output value. And this field is divided into supervised learning, which is predicts the value of input data and is classified with the given label, and unsupervised learning that is performed in a state where data is not labeled and is often called clustering process.

In this way, to distinguish new examples, the AI strategies based on supervised learning and unsupervised learning and profound using counterfeit neural systems have been effectively examined, and that is to look at the effectiveness of recognizing monetary misrepresentation exchanges. The detection process of the financial fraud contains sampling process for class imbalance problem and feature selection process for an accurate model.

The Banking Industry:

By the applying fraud detection techniques, banks will offer outstanding types of assistance; it is a system of billion gadgets associated through the web that is used by individuals in everyday occasions will make banks gain many competitive advantages, and it will be the most intelligent system. Furthermore, it will foresee the client's needs through information gathered and examined, at that point gives arrangements that can assist clients with taking smart financial choices [18].

Recently, the analytics insights and the statistical modeling is considered to be an important element in the industry to improve the optimization, forecast and the decision process of the operation. Nonetheless, the banks and the financial establishments focus on income progress, the higher borders through its operational efficiency, particularly better hazard the executives, and improved client closeness. As a result, the development and create arrangement and programming designs need to manage these difficulties, with the need of contemplating duration and expenses, that make them at the forefront of the markets [18].

In short, the information becomes essential to identify fake conduct that is the reason Web of Things collects enormous information from different gadgets and offer them through the web to forestall any fake activities.

IV. COMPUTATIONAL INTELLIGENECE FOR FRAUD DETECTION

We surveyed the most recent methods to identify irregularity and trust handing-off in the IoT condition. Additionally, we concentrated on inspecting the strategies and calculations. To forestall budgetary extortion, they performed investigation through numerical simulations by utilizing a guarded system and introduced a contextual analysis on the discovery of Fake or false news sources in the social Environment of The IoT condition [9]. Additionally, there are different approaches to identify misrepresentation, which is an AI that uses as a web administration based community plot for MasterCard extortion discovery. The fraud detection system isolated into two phases; in the principal stage, looking at the approaching exchange against the exchange history to recognize the abnormality utilizing the Boat calculation, and adaptable calculation. The false alert rate presumed peculiarities are checked with the extortion history database and conclude that the distinguished irregularities. As a result, the deceitful exchange or any momentary change is going through profile in the second stage [10] BLAST and SSAHA algorithms are a productive strategy to look at the spending conduct of clients [11], to compute the likelihood from the client's current monetary data, and to construct a multilayer model of program conduct. Syeda et al. in (2002) they proposed fluffy neural systems “fuzzy” that run on equal machines to accelerate the standard creation for credit, in this strategy Method, the Granular Neural Network (GNN) technique that utilizes a fuzzy neural system which depends on information / knowledge discovery [12]. And there are other methods of fraud detection like a combined method of neural networks, logistic regression, and the decision tree [13], and also the Self-Organizing Map Neural Network which known briefly as (SOMNN).

Neural Networks

Due to Aleskerovet al. “Developing a system, for the data extraction, based on the neural network to detect fraud in the credit cards. And also, the (CARD WATCH) proposed system contains three layers of automatic binding structures. And they used a set of structured data to train and test the system. Consequently, the results were showing high successful fraud detection rates.” In, P-RCE's the neural network was applied to detect the fraud in the credit card: P-RCE is a type of the radial base function networks that are applied commonly to the tasks of pattern recognition. Krinker et al. is another model that was proposed to detect fraud in real-time, and it based on the bidirectional neural networks. In this method, they used telephone transaction data of a wide range provided by the credit card company. And it supposed that the system outperforms the algorithm based on the rule in terms of the false-positive rate. In a parallel Granular Neural Network (GNN) is proposed to speed up the data mining and the knowledge discovering to detect the fraud in the credit cards. And it is a type of indefinite neural network which is based on knowledge discovery (FNNKD). So, the basic data set was extracted from the SQL Server database containing Visa Card transactions and then pre-processed for fraud detection. As a result, they received less training errors on average in the presence of a larger training data set.

Support Vector Machines

The Support Vector Machines (SVM) is a procedure of measurable learning and has application fruitful in issues scope. Also, the first presentation by Cortes and Vapnik (1995), and it has been showed to be useful in an assortment of grouping errands.

By Y. Sahin and E. Duman, Their Research “Detecting Credit Card Fraud by Decision Trees and Support Vector Machines”[] they made data Distribution with a respect to highly Imbalanced classes. Due to them the time period that used to build their sample which Included about 978 fraudulent

records and about 22 million normal ones with a ratio of about 1:22500.t

Decision Tree

The learning system concept was introduced, then the decision tree method was developed, the 'C4.5 technique' which could handle continuous data by (Quinlan, 1993) and another technique "ID3 method" by (Quinlan, 1986),. the decision tree is a table of tree shape with lines attached to the available nodes that be. Eeach node is a sub node which is followed by only one node that is assigned by the classification or more nodes.

. Due to, (van et al., 2001), the similarity trees which yielded proven results worked on trees decision especially on the other type of the fraud, inductive decision tree in order to create a system of the intrusion detection.

Self-Organizing Map Neural Network

The self-organized map, it is an architecture which suggested for the artificial neural networks, was explained by the presenting of simulation experiments and practical applications. The self-organizing map has the property of effectively creating spatially organized internal representations of various features of input signals and their abstractions. One result of this is that the self-organization process can discover semantic relationships in sentences. Brain maps, semantic maps, and early work on competitive learning are reviewed. The self-organizing map algorithm (an algorithm which order responses spatially) is reviewed, focusing on best matching cell selection and adaptation of the weight vectors. Suggestions for applying the self-organizing map algorithm, demonstrations of the ordering process, and an example of hierarchical clustering of data are presented. Fine tuning the map by learning vector quantization is addressed. The use of self-organized maps in practical speech recognition and a simulation experiment on semantic mapping are discussed

V. RESULTS

As a result, the notable the well-known machine learning method has a higher fraud detection rate than the other artificial neural network AI technique has a higher misrepresentation recognition rate than the counterfeit neural system. Due to [17] the average discovery rate was 0.98618 when the entirety of the calculations algorithms were used.

In the AI model of machine learning, the trials were performed from clustering systems, for example, EM, straightforward, XMeans, Farthest First, and Density-Based calculations in grouping calculations, EM calculation arrived at a normal of 0.99862 in extortion discovery. The density-Based calculation positioned in the second Hight in misrepresentation discovery and arrived at 0.98788. In characterization calculations, Relapse arrived at an average of 0.99971 in misrepresentation identification. Also, Random Forest reached an average of 0.99969 and the second top in fraud C4.5 positioned the third level by arriving at 0.99943.[17]

VI. CHALLENGES

The monetary exchange contains an information imbalanced issue that could misguide the detection process, according to this problem, the exploration creates different datasets utilizing Manufactured Minority over Sampling Technique (SMOTE) and Random under Sampling (RUS) for the more precise trial. The SMOTE is an oversampling strategy that utilizes a technique for producing subjective models as opposed to just oversampling through substitution or duplication [14], however, RUS was applied for scaling down the ordinary exchanges by separating test information randomly for the class awkwardness issue.

The Recent current design utilized in banks doesn't reach yet the degree of affectability required, because of the large size of data that needs to be prepared and contrasted and defined due to banking rules. The existing of architecture does not mean that the financial institution will become more efficient to identify extortion, because of numerous reasons that we refer to underneath with current design as well:

- None of the frameworks utilize the IoT innovation to gather information.
- The needing utilization of ETLs for Concentrate, change, and burden capacities.
- The losing of unique crude information for any new data after processing.
- Limitation as far as handling a typical pool of capacity information.

VII. OPPORTUNITIES

Feature selction. IT has been demonstrated to be efficient for AI issues. The goals of Feature selection incorporate structure less difficult and progressively understandable models, and furthermore it incorporates planning to evacuate to reasonable information [15]. Feature selection can be separated into covering and channel techniques. The covering strategy relied on the predictive presentation of a specific learning calculation to assess the selected features.

The disadvantage of the wrapper technique is that the inquiry space could be immense and it is generally more costly than different strategies. The filtering strategy is free of any learning calculations and depends on specific attributes of information to survey the significance of features. And these characteristics are inspected dependent on the scores as per the assessment models, and the most minimal scored highlights and features are evacuated [16]. For this reason, the filter-based feature selection algorithms is for the element determination technique, which is the quickest and furthermore reasonable for efficient use. It is suggested that the feature determination technique that can be applied to the automation system; first, the system model process is constructed by applying the element determination technique on the unsupervised learning calculation, and

afterward applied the administered learning calculation later for exact analysis dependent on the above exploratory outcomes by open dataset and genuine dataset, at the end compared the last precision of the proposed AI-based discovery model and the recognition exactness of models utilizing fake profound neural systems. What's more, utilizing Huge Information advances in this sort of an administration can roll out tremendous improvements in the current adaptation of handling in execution points of view, which is considered as a basic viewpoint and most significant component continuously extortion observing logic because each millisecond turns out to be essential to forestall financial influence on the banking institutions.

VIII. CONCLUSION AND FUTURE WORKS

So, the process depends on the AI technique that contains the component determination process, which dependent on the filter strategy, the clustering method, and the classification process. Furthermore, the test results show that the two featured selection algorithms which are filtered algorithm and ranked algorithm AI techniques has higher location proficiency than neural systems, and the neural systems arrived at an especially high discovery exactness at 95: 1 and 99: 1 ratio[17], which is about like the genuine proportion in reality. In future work, we intend to improve the precision and handling time of the monetary misrepresentation process continuously joined with both the AI-based procedure and profound artificial neural systems.

All these methods and techniques are used to recognize fake cases before it occurs or during the handling of any online exchange that meets any misrepresentation decide that defined by the bank's risk department in the database. In our design, we concentrated on MapReduce, one of the fundamental empowering approaches for satisfying expanded Misrepresentation frameworks needs by utilizing high equal preparation, information storage, investigation, and web-based handling on countless items.

REFERENCES

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

[2] Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B.: APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. Decis. Support Syst. 75, 38–48 (2015)

[3] Duman, E., Buyukkaya, A., Elikucuk, I.: A novel and successful credit card fraud detection system implemented in a Turkish bank. In: 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW). IEEE (2013)

[4] Bahnsen, A.C., et al.: Cost sensitive credit card fraud detection using Bayes minimum risk. In: 2013 12th International Conference on Machine Learning and Applications (ICMLA), Vol. 1. IEEE (2013)

[5] Wei, W., et al.: Effective detection of sophisticated online banking fraud on extremely imbalanced data. World Wide Web 16(4), 449–475 (2013)

[6] k. corp, “Mobile payments fraud survey report,” 2016.

[7] “Javelin strategy and research,” 2016.

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

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

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

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

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

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

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

[15] K. E. P. Baksai, Feature Selection to Detect Patterns in Supervised and Semi Supervised Scenarios, Ph.D. thesis, Pontificia Univer- sidad Cato´lica de Chile, 2010.

[16] L. Jundong et al., “Feature selection: A data perspective,” ACM Computing Surveys (CSUR), vol. 94, 2017.

[17] Dahee Choi and Kyungho Lee “An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation” Security and Communication NetworksVolume 2018, Article ID 5483472, 2018.

[18] Abdeljalil Boumlik, Mohamed Bahaj. "Chapter 35 Big Data and IoT: A Prime Opportunity for Banking Industry", Springer Science and Business Media LLC, 2018.