



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Novel Technique of Anti-Money Laundering Using Deep Learning

Sharanpreet Kaur

Research Scholar

Department of Computer Science and Engineering

Desh Bhagat University

Navneet Kaur Sandhu

Assistant Professor

Department of Computer Science and Engineering

Desh Bhagat University

Abstract: With the advancement of technology, the potential risk of Money laundering is also increased significantly; Advance deep learning techniques with availability of big data can be seen as a promising solution to protect money laundering. In this paper we have proposed a deep learning based anti-money laundering algorithm. The algorithm is being tested on publically available data-set. And the performance of proposed algorithm has shown the state of art accuracy. The performance of proposed deep learning algorithm has been compared with other Machine learning algorithm and proposed deep learning algorithm outperformed the rest of the algorithms.

Keywords: Deep Learning.

Introduction: Assessing the impact of ML, being legally a crime but that seems to have no victims as it cannot be seen, the ruin it causes can be destructive to the financial sector, especially in the case of a developing country. Today, the economy of the country depends enormously upon the growth of technology. It eases the job but it's more challenging. The IMF estimated the total laundered money worldwide was 2% to 5% of global GDP in 1998 . The growth of the ML problem is broad, and a huge amount of this crime goes undeclared, thus hidden. Authorities around the world have been fighting to get hold of the whole problem. Based on implementations of AML policies and legal sectors, developments are made but they fall short of investigation on the consequences and outcomes of ML in the development economies. Hence, no adequate data is available to reach an accurate conclusion. The goal is allowing the use of money gained from carrying out activities illegally while concealing the origin from government authorities.

Literature survey: Le Khac et al. [1] Recently, money laundering is becoming more and more sophisticated, it seems to have moved from the personal gain to the cliché of drug trafficking and financing terrorism. This criminal activity poses a

serious threat not only to financial institutions but also to the nation. Today, most international financial institutions have been implementing anti-money laundering solutions but traditional investigative techniques consume numerous man-hours. Besides, most of the existing commercial solutions are based on statistics such as means and standard deviations and therefore are not efficient enough, especially for detecting suspicious cases in investment activities. In this paper, we present a case study of applying a knowledge-based solution that combines data mining and natural computing techniques to detect money laundering patterns. This solution is a part of a collaboration project between our research group and an international investment bank. Liu, Rui, et al. [2] This paper presents a core decision tree algorithm to identify money laundering activities. The clustering algorithm is the combination of BIRCH and K-means. In this method, decision tree of data mining technology is applied to anti-money-laundering filed after research of money laundering features. We select an appropriate identifying strategy to discover typical money laundering patterns and money laundering rules. Consequently, with the core decision tree algorithm, we can identify abnormal transaction data more effectively. Tang, Jun. [3] Statistical learning theory (SLT) is introduced to improve the embarrassments of anti-money laundering (AML) intelligence collection. A set of unusual behavior detection algorithm is presented in this paper based on support vector machine (SVM) in order to take the place of traditional predefined-rule suspicious transaction data filtering system. It could efficiently surmount the worst forms of suspicious data analyzing and reporting mechanism among bank branches including enormous data volume, dimensionality disorder with massive variances and feature overload. Luo, Xingrong. [4] Money laundering activities in financial markets have been increasingly serious recent years. Although efforts on anti-money activities started at an early stage, the solutions seem to be restricted to a strategic level. Besides, even some research pioneers at employing data mining techniques to anti-money laundering, the situation in China is still difficult. To this end, in this paper, after presenting the systematic view of the data mining framework of anti-money laundering, we propose a classification based algorithm to effectively

detect suspicious transactions. Specifically, we consider the financial transactions as a data stream, and try to construct a classifier based on a set of mined frequent rules. Our experiments on a simulated transaction dataset based on real world banking activities prove the efficiency of our proposed method. Keyan et al. [5] The selection of parameters of SVM model will affect the identification effect of suspicious financial transactions, this paper proposes the cross validation method to find the optimal SVM classifier parameters to solve this problem. Cross validation method finds the optimal parameters based on the highest classification accuracy rate through grid search, it can effectively avoid the state of over-learning and less learning, and greatly improves the overall performance of the classifier.

Liu, Xuan et al. [6] Developing effective suspicious activity detection methods has become an increasingly critical problem for governments and financial institutions in their efforts to fight money laundering. Previous anti-money laundering (AML) systems were mostly rule-based systems which suffered from low efficiency and could be easily learned and evaded by money launders. Recently researchers have begun to use machine learning methods to solve the suspicious activity detection problem. However nearly all these methods focus on detecting suspicious activities on accounts or individual level. In this paper we propose a sequence matching based algorithm to identify suspicious sequences in transactions. Our method aims to pick out suspicious transaction sequences using two kinds of information as reference sequences: 1) individual account's transaction history and 2) transaction information from other accounts in a peer group. By introducing the reference sequences, we can combat those who want to evade regulations by simply learning and adapting reporting criteria, and easily detect suspicious patterns. The initial results show that our approach is highly accurate. Singh et al [7] Annually, money laundering activities threaten the global economy. Proceeds of these activities may be used to fund further criminal activities and to undermine the integrity of financial systems worldwide. For these reasons, money laundering is recognized as a critical risk in many countries. There is an emerging interest from both researchers and practitioners concerning the use of software tools to enhance detection of money laundering activities. In the current economic environment, regulators struggle to stay ahead of the latest scam, and financial institutions are challenged to ensure that they can identify and stop criminal activities, while ensuring that legitimate customers are served more effectively and efficiently. Effective technological solutions are an essential element in the fight against money laundering. Improved data and analytics are key in assisting investigators to focus on suspicious activities. Continually evolving regulations, together with recent instances of money laundering violations by some of the largest financial institutions, have highlighted the need for better technology in managing anti-money laundering activities. This study explores the use of visualization techniques that may assist in efficient identification of patterns of money laundering activities. It demonstrates how link analysis may be applied in detecting suspicious bank transactions. A prototype application (AML2ink) is

used for proof-of-concept purposes. Umadevi et al. [8] Banks are faced with the challenge of adapting to the ever-changing risk and methods related to money laundering. Money laundering is a complex, dynamic and distributed process. Some Anti-Money Laundering (AML) Systems simply transforms vast quantities of data into vast numbers of reports that do not facilitate timely detection or effective interdiction. Transaction Flow Analysis (TFA) system is proposed to get over this issue. The main parts of this TFA system is, first, bank transaction importer, which is not bound to any file format. Second, application of frequent pattern mining and transaction mining algorithms to detect money laundering, which implements distributive box and collective box. Third, clustering the transaction and detecting suspicious clusters by recognizing laundering methods and roles of the offender. Finally obtained clusters and frequent patterns can be visualized in schema and timeline diagrams. It allows the bank or police analyst to find suspected money flows and suspicious transactions. Colladon et al. [9] This research explores the opportunities for the application of network analytic techniques to prevent money laundering. We worked on real world data by analyzing the central database of a factoring company, mainly operating in Italy, over a period of 19 months. This database contained the financial operations linked to the factoring business, together with other useful information about the company clients. We propose a new approach to sort and map relational data and present predictive models – based on network metrics – to assess risk profiles of clients involved in the factoring business. We find that risk profiles can be predicted by using social network metrics. In our dataset, the most dangerous social actors deal with bigger or more frequent financial operations; they are more peripheral in the transactions network; they mediate transactions across different economic sectors and operate in riskier countries or Italian regions. Finally, to spot potential clusters of criminals, we propose a visual analysis of the tacit links existing among different companies who share the same owner or representative. Our findings show the importance of using a network-based approach when looking for suspicious financial operations and potential criminals.

Luo et al. [10] Money laundering activities in financial markets have been increasingly serious recent years. Although efforts on anti-money activities started at an early stage, the solutions seem to be restricted to a strategic level. Besides, even some research pioneers at employing data mining techniques to anti-money laundering, the situation in China is still difficult. To this end, in this paper, after presenting the systematic view of the data mining framework of anti-money laundering, we propose a classification based algorithm to effectively detect suspicious transactions. Specifically, we consider the financial transactions as a data stream, and try to construct a classifier based on a set of mined frequent rules. Our experiments on a simulated transaction dataset based on real world banking activities prove the efficiency of our proposed method. Dreżewski et al.

[11] Criminal analysis is a very complex task requiring to process huge amounts of data coming from different

sources such as billings and bank account transactions in order to gain knowledge useful for an investigator. In order to support human analytic capabilities, dedicated software tools are needed, and therefore Money Laundering Detection System (MLDS) was proposed as one of such tools in our previous paper. In this paper, the social network analysis component for this system is presented. The component makes it possible to use data from bank statements and the National Court Register and construct and analyze social networks during an investigation into money laundering cases. The system can assign roles to persons from the network and allows for analysis of connections between them. The paper also includes results of experiments aimed at investigating the performance of the implemented algorithms and the correctness of the analysis. Chao et al. [12] Trade-based Money Laundering, a new form of money laundering using international trade as a signboard, always appears along with speculative capital movement which has been accepted as the most concerned and consensus incentive giving rise to the collapse of the financial market. Unfortunately, preventing money laundering is very difficult since money laundering always has a plausible trade characterization. To reach this goal, supervision for regulator and financial institutions aims to effectively monitor micro entities' behavior in financial markets. The main purpose of this paper is to establish a monitoring method including accurate recognition and classified supervision for Trade-based Money Laundering by means of knowledge-driven multi-class classification algorithms associated with macro and micro prudential regulation, such that the model can forecast the predicted class from the concerned management areas. Based on empirical data from China, we demonstrate the application and explain how the monitor method can help to improve management efficiency in the financial market. Wang et al. [13] Money laundering (ML) involves moving illicit funds, which may be linked to drug trafficking or organized crime, through a series of transactions or accounts to disguise origin or ownership. China is facing severe challenge on money laundering with an estimated 200 billion RMB laundered annually. Decision tree method is used in this paper to create the determination rules of the money laundering risk by customer profiles of a commercial bank in China. A sample of twenty-eight customers with four attributes is used to induced and validate a decision tree method. The result indicates the effectiveness of decision tree in generating AML rules from companies' customer profiles. The anti-money laundering system in small and middle commercial bank in China is highly needed. Luo, Xingrong. [14] Money laundering activities in financial markets have been increasingly serious recent years. Although efforts on anti-money activities started at an early stage, the solutions seem to be restricted to a strategic level. Besides, even some research pioneers at employing data mining techniques to anti-money laundering, the situation in China is still difficult. To this end, in this paper, after presenting the systematic view of the data mining framework of anti-money laundering, we propose a classification based algorithm to effectively detect suspicious transactions. Specifically, we consider the

financial transactions as a data stream, and try to construct a classifier based on a set of mined frequent rules. Our experiments on a simulated transaction dataset based on real world banking activities prove the efficiency of our proposed method. Le Khac, Nhien An, et al. [15] Today, money laundering (ML) poses a serious threat not only to financial institutions but also to the nation. This criminal activity is becoming more and more sophisticated and seems to have moved from the cliché of drug trafficking to financing terrorism and surely not forgetting personal gain. Most of the financial institutions internationally have been implementing anti-money laundering solutions (AML) to fight investment fraud activities. However, traditional investigative techniques consume numerous man-hours. Recently, data mining approaches have been developed and are considered as well-suited techniques for detecting ML activities. Within the scope of a collaboration project on developing a new data mining solution for AML Units in an international investment bank in Ireland, we survey recent data mining approaches for AML. In this paper, we present not only these approaches but also give an overview on the important factors in building data mining solutions for AML activities.

Salehi et al. [16] Money Laundering is the process of creating the appearance that large amounts of money obtained from serious crimes, such as drug trafficking or terrorist activity, originated from a legitimate source. Through money laundering, the launderer transforms the monetary proceeds derived from criminal activity into funds with an apparently legal source. The system that works against Money laundering is Anti-Money Laundering (AML) system. The existing system for AntiMoney Laundering accepts the bulk of data and converts it to large volumes reports that are tedious and topsy-turvy for a person to read without any help of software. To develop a structure to research in datamining, we create a taxonomy that combines research on patterns of observed fraud schemes with an appreciation of areas that benefit from a productive application of data mining. The aim of this study was to review research conducted in the field of fraud detection with an emphasis on detecting honey laundering and examine deficiencies based on data mining techniques. Which include a set of predefined rules and threshold values. In addition to this approach, data mining techniques are very convenient to detect money laundering patterns and detect unusual behavior. Therefore, unsupervised data mining technique will be more effective to detect new patterns of money laundering and can be crucial to enhance learning models based on classification methods. Of course, the development of new methods will be very useful to increase the accuracy of performance. Gao, Zengan. [17] Financial institutions' capability in recognizing suspicious money laundering transactional behavioral patterns (SMLTBPs) is critical to antimoney laundering. Combining distance-based unsupervised clustering and local outlier detection, this paper designs a new cluster based local outlier factor (CBLOF) algorithm to identify SMLTBPs and use authentic and synthetic data experimentally to test its applicability and effectiveness. Yang et al.[18] Online payment becomes a convenient way to launder money with

the development of e-commerce. In order to solve this problem, we constructed an anti-money laundering system as a service function of union-bank centre. This system can monitor and analyze the transaction data dynamically, and provide auxiliary judgment and the decision support for anti-money laundering. It utilized many technologies synthetically, such as multi-agent neural network, text mining, genetic algorithms, velocity analysis and case-based reasoning technologies to find out the hidden money laundering behavior. The logical frame of this system was proposed and some of the key technologies were introduced. Finally, we discussed the practicability and the development of this system and the union-bank online payment mode. In our opinion, this mode will become the lead online payment mode and its anti-money laundering service system will perform a very useful role.

Proposed Methodology:

Presently various algorithms are available for clustering the proposed data. In in the existing work, they used K mean clustering, Support Vector Machine and ANN claimed to achieve optimum solutions. But still we can see that there is vast scope of improvement in AML (Anti-Money Laundering).

As more and more data is getting generated these days, there is need of improvement in the existing algorithm which can perform well with new data and can be scaled onto real time applications.

Our derived solution is a classification system which is based upon Deep Learning. The data-set we have taken is open source at Kaggle website and used with the aim of achieving state of art accuracy.

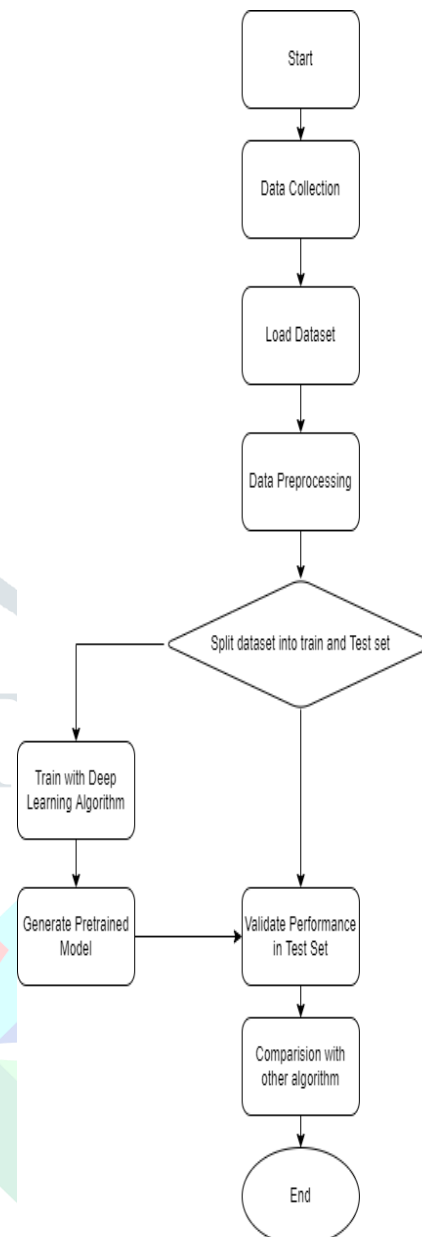


Figure 1 Flowchart of Proposed System

Data Collection: For data collection, we have started web scrawling to reach out to relevant open source datasets. From the beginning, we were looking for type of authenticated datasets with aforementioned labels so that we can train the classification algorithm. The data-set we have taken is open source at Kaggle website. Every attribute of dataset is taken from real-time activities with mentioned labels. One of the major challenges that we faced during the data gathering process was the unbalanced dataset. An unbalanced dataset can result in generating adulterated, bad quality or inaccurate results. Since most of the activities in real world are genuine, very few of them can actually be used for money laundering.

Load Data-set: This is relatively the most important step in our entire algorithm. It is because the data-set we have holds very limited feature based information that makes us unable to apply feature extraction and selection method on it. So if we push the data-set without filtering, Machine learning model can present the problem of over fitting towards the authenticated transaction type. Over fitting is such a stage in Machine Learning when model cannot be generalized into more datasets.

Data Preprocessing: In the collected data-set, very high records of genuine transactions are found and out of them, very few are related to money laundering. To balance the data-set, we have applied thresholding that enables us to remove those transactions in which very small amounts were involved. All of these small transactions with respect to the amount of transactions are genuine. Keeping the big transaction amount into consideration, we move forward with those transactions. After filtering the small amount based transactions of data-set, we are left with balanced data-set where fraud and genuine transactions are in balance.

Split data-set into Train and test set: Train/Test is a method to measure the accuracy of your model. It is called Train/Test because you split the the data set into two sets: a training set and a testing set. 80% for training, and 20% for testing. We train the model using the training set.

Implementation of Algorithm: The implementation of our algorithm i.e. Deep-Learning. We have used Deep-Learning to generate accurate decisions regarding detection of money laundering activities. The algorithm helps in detecting any suspicious or fraudulent activities that might threat or harm the organizations. In terms of continuous transactional data in money laundering case, this algorithm works well on processing the data and leading us to accurate results. That's why we have used Deep-Learning instead of using other algorithms to derive the exact results.

Result and discussion:

For the simulation purpose, we used Google Colab that is powered with GPU and available free of cost. The configuration of our virtual machine at Google Colab is 12 GB RAM, with Linux Kernel and VRAM of 12 GB. The entire implementation has been done by us in Python 3.8.

The Table1 below represents the results of the first sets of training data that were available. Each of the dataset contained 10,000 records and different sets of transaction. Those transactions are then run through the algorithms and the following results are observed.

	C4.5 Accuracy	ANN Accuracy	Proposed Deep-Learning Model
Accuracy	92.08%	93.93%	98.68%
Precision	92.29%	94%	98.69%
Recall	92.68%	95.21%	98.20%
F-Measure	92.34%	95.26%	98.47%

Table1 Showing the comparison of accuracy of Proposed algorithm with the existing algorithms.

In table 1, an intensive comparison has been given with all algorithms, and presents the comparison of accuracies derived using different algorithms that are C4.5, ANN, Proposed Algorithm. Implementation of Proposed Algorithm leads to best results in terms of accuracy i.e. 98.68%. This table shows that the Proposed Algorithm is

the most suitable for eliminating the money laundering activities.

As Shown in the table Proposed Algorithm outperformed other algorithms well and succeeded in curbing Money Laundering and related fraudulent activities. Proposed Deep-Learning model is perfect to be worked on especially when it comes to detecting money laundering.

Conclusion:

In this paper, we have covered all the theories related to the Anti-Money Laundering. We aimed for finding a good state of the art algorithm for the purpose of Anti-Money Laundering. we have made the best use of Machine Learning and the algorithms to derive the best possible results in fighting against Money Laundering or any similar fraudulent activities. Proposed algorithm came up with the maximum accuracy in detecting money laundering activities and outperformed other algorithms in the same. The Proposed algorithm combating the money laundering with the accuracy of 98.68%.

References:

1. Le Khac, Nhien An, and M-Tahar Kechadi. "Application of data mining for anti-money laundering detection: A case study." 2010 IEEE International Conference on Data Mining Workshops. IEEE, 2010.
2. Liu, Rui, et al. "Research on anti-money laundering based on core decision tree algorithm." 2011 Chinese Control and Decision Conference (CCDC). IEEE, 2011.
3. Tang, Jun, and Jian Yin. "Developing an intelligent data discriminating system of anti-money laundering based on SVM." 2005 International conference on machine learning and cybernetics. Vol. 6. IEEE, 2005.
4. Luo, Xingrong. "Suspicious transaction detection for anti-money laundering." *International Journal of Security and Its Applications* 8.2 (2014): 157-166.
5. Keyan, Liu, and Yu Tingting. "An improved support-vector network model for anti-money laundering." 2011 Fifth International Conference on Management of e-Commerce and e-Government. IEEE, 2011.
6. Liu, Xuan, Pengzhu Zhang, and Dajun Zeng. "Sequence matching for suspicious activity detection in anti-money laundering." *International Conference on Intelligence and Security Informatics*. Springer, Berlin, Heidelberg, 2008.
7. Singh, Kishore, and Peter Best. "Anti-money laundering: using data visualization to identify suspicious activity." *International Journal of Accounting Information Systems* 34 (2019): 100418.

8. Umadevi, P., and E. Divya. "Money laundering detection using TFA system." (2012): 17-17.
9. Colladon, Andrea Fronzetti, and Elisa Remondi. "Using social network analysis to prevent money laundering." *Expert Systems with Applications* 67 (2017): 49-58.
10. Luo, Xingrong. "Suspicious transaction detection for anti-money laundering." *International Journal of Security and Its Applications* 8.2 (2014): 157-166.
11. Dreżewski, Rafał, Jan Sepielak, and Wojciech Filipkowski. "The application of social network analysis algorithms in a system supporting money laundering detection." *Information Sciences* 295 (2015): 18-32.
12. Chao, Xiangrui, et al. "Behavior monitoring methods for trade-based money laundering integrating macro and micro prudential regulation: a case from China." *Technological and Economic Development of Economy* 25.6 (2019): 1081-1096.
13. Wang, Su-Nan, and Jian-Gang Yang. "A money laundering risk evaluation method based on decision tree." *2007 International Conference on Machine Learning and Cybernetics*. Vol. 1. IEEE, 2007.
14. Luo, Xingrong. "Suspicious transaction detection for anti-money laundering." *International Journal of Security and Its Applications* 8.2 (2014): 157-166.
15. Le Khac, Nhien An, et al. "An investigation into data mining approaches for anti-money laundering." *Proceedings of International Conference on Computer Engineering and Applications (ICCEA 2009)*. 2009.
16. Salehi, Ahmad, Mehdi Ghazanfari, and Mohammed Fathian. "Data mining techniques for anti-money laundering." *International Journal of Applied Engineering Research* 12.20 (2017): 10084-10094.
17. Gao, Zengan. "Application of cluster-based local outlier factor algorithm in anti-money laundering." *2009 International Conference on Management and Service Science*. IEEE, 2009.
18. Yang, Qifeng, Bin Feng, and Ping Song. "Study on anti-money laundering service system of online payment based on union-bank mode." *2007 International Conference on Wireless Communications, Networking and Mobile Computing*. IEEE, 2007.