Developing Anti Money Laundering Identification using Machine Learning Techniques

Kidist Sintayehu[1] & Hussien Seid (PhD)[2]

[1]*Artificial Intelligence and Robotics Center of Excellence, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia.* [2]*Department of Software Engineering, College of Electrical and Mechanical Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia.*

ABSTRACT

There are several areas that can be applied the machine learning techniques, one of which is the finance and economics sector, it can effectively help and assist on decision support on fraud investigation. In this sector the skills and methods of fraud is becoming more complicated. In financial sector there are a lot of fraud mechanisms from them money laundering is crucial activity to fraud. The main purposes of this study develop and recommend a machine learning algorithm for identification of Anti-Money Laundering. The bank data set collected from Commercial Bank of Ethiopia which has 14740 with number of missing value and then after effectively pre-processing the data set, the five different types of machine learning algorithms were used, these are support vector machine, Logistic Regression, Naïve Bayes Decision tree, and Random forest. From those machine learning algorithms, Random Forest the best performance with 99.1% accuracy.

Keywords: Bank transaction; Fraud; Money laundering; Financial services; Machine learning.

1. Introduction

Money laundering is the process of changing large amount of money from illegal activities and crimes into legal sources [1]. It is a crime in many regions having different kinds of meanings. Organized crime and the accumulation of unknown resources are dangerous activities [2],[3]. Money laundering is not exactly known. But majority of financial believes, it began in the 1960's and was carried out under various circumstances. Hiding money from various illegal activities by properly hiding the sources of money and then depositing it in various accounts and using it for illegal purposes [4],[5]. On the other hand, investing in the highest quality goods and services is a legitimate use. Funds raised from money laundering are used to directly and indirectly support and assist in activities such as terrorism. Fraudulent individuals use a variety of methods to defraud money. The first of these involves storing and storing the money collected in a trustworthy place or person. Next the hidden money, it is an attempt to destroy the source after that the swindler involves spending money in a very complex transaction system. Finally, after the transaction activity is over, the fraudster indirectly or unknowingly returns the money to himself in a way that the investigator does not understand. Process of money laundering can be defined as the following, the first one is placement, layering and integration.

According to UN and Financial Accountability Transparency & Integrity (FACTI) Panel Interim reports [6], the world loses more than one trillion American dollar by this money laundering which is 2.7% of the Global GDP. And this this number of might be increases by double after 5-10 years. There are different mechanisms to fight against money laundering improve and standardize systems with technologies, know the customer that work with the financial company, use some method of data analytics to get the form and train and update the worker. The economic factors of money laundering are includes performing this illegal activity behind the scenes while entering the real and legitimate private businesses sectors and may have erase good name of this companies, can affect legitimating of the economic rules. On the other side it can also affect the social aspects by increasing terrorism

organization, drug dealers, smugglers and other criminal activities. If this money laundering becomes more complex then can control the system of government of the environment.

Anti-money laundering is set of rules and regulations that can be used to fight against money laundering in financial system. In the financial sector, various issues are happening around the world, Anti-money laundering is a critical part of financial anti-crime programs. The worldwide environment of money laundering needs global standards and in order to reduce the ability of criminals to launder their proceeds and carry out criminal activities. Anti-money laundering activities are designed to prevent the ability of criminals to use their illegal gains [7]. Using artificial intelligence technologies to design a modern to increase the security related issued to the financial system modification. This study aimed to design a model that can be used to identify this money laundering and alert the financial institution using machine learning techniques. This techniques used to classify the transaction in the stage of placement before goes to the final step which is difficult to investigate the property after integration investing in/outside the country. Supervised machine learning technique for identifying and alert Money Laundering hidden patterns, groups and transactions in finance.

2. Statement of the Problem

Money laundering has a huge impact on Banks and other financial sectors. According to the research done from 2016 to 2021, between ten million dollars and two billion dollars will be lost every year in this Money boundary activities. In addition, more than twenty million dollars will be taken out of the country. To overcome the various activities of money laundering, banking sectors should have a well-organized anti-money identification system. The common technique that Bank sectors follow to identify fraudulent activity is Manual with rule-based. A rule-based technique requires an extensive list of specifications. This makes it very difficult to identify all the fraudsters and list them in time. Money Laundering is the process of hiding the source of money from illegal activities into legal sources. Money Laundering is the process of hiding the source of money from illegal activities into legal sources, which has a serious impact on the country's economy. Manual and rule-based money laundering identification technique can easily be bypassed. Machine learning has proven to be effective in detecting Money Laundering, but there is a trade-off between detection accuracy. Since money laundering activities are affecting Bank sectors it should be detected before making huge damage.

Since money laundering activities are affecting Bank sectors, it has to be detected before making huge damage. Generally, Using machine learning, Anti-money laundering mechanisms help in the protection of the financial systems, and institutions to reduce the risks of suspected types of transactions and other human and machine errors [8]. New technologies such as changing the system fully or partially by an automated system can speed up and reduce mistakes by giving accurate results, which minimizes a gap between fraud and policies.

3. Research Objective

The general objective of this study is to develop an anti-money laundering identifyion and identification model using machine learning techniques.

The specific objectives of this study are as follows:

• A comprehensive Literature review will be conducted to learn the current state of the technology.

- Classify and identify the types of fraud that are found in the bank transaction system.

- Study and identify the limitation and gaps of traditional fraud detection methods and techniques.

- To prepare the machine learning models.

- Develop a model for anti-money laundering detection using machine learning.

- Test and evaluate the proposed model by test dataset.

4. Literature Review

Financial institutions have importance to a country's economic growth is enormous. However, adding sustainable investment to the economy will enable us to sustain sustainable growth. One of the things that contribute to the sustainability of the economy is adding technological assets. The bank is one of the financial institutions that need high technological advancement. The banking system plays an important role in the modern economic world. Banks collect the savings of the individuals and lend them out to business- people and manufacturers. The banking system can create money. There are currently more than 18 public and private banks in Ethiopia. Of these, the Commercial Bank of Ethiopia has the largest wealth and customer base. In these banks, there are many transactions per day. The fact that such many transactions in one country show that there is an active economic activity in one country. While this transfer is good for a country's economic growth, it can also be done legally and illegally.

Although companies use a variety of methods to prevent illegal money transactions, they are more likely to be involved in illegal money transactions because they do not use the latest technology. However, it is important to try to reduce these obvious illegal activities through the use of modern data science techniques. Machine learning is a part of artificial intelligence that can solve complex problems like fraud detection, malware detection, image recognition, virtual reality, and so on. Machine learning is a part of artificial intelligence that can solve complex problems like fraud detection, malware detection, image recognition, virtual reality, and so on. In machine learning model of the system learn from a given data; input and output. And create decisions with less human interference. Therefore for this proposed problem Machine learning has a solution with a great opportunity. There are different types of banks including Commercial banks, Retail banks, Investment banks, Insurance companies, Internet banking and saving and loan association. Money laundering is a process of illegally making huge amounts of money that came from different criminal activities, for this illegal money to appear as legal source of money. It is by another simple explanation washing the dirty money through legal financial institutions such as bank appear as clean money. Money laundering has a lot of impact on economy, social and politically of country. This addressed problem is also affecting the whole world of money exchange system.

Money laundering have three stages to perform illegal activities. The first stage is placement. In this stage the process of depositing cash into banks using multi-small transaction. First large amount of money broken in to different accounts to transferred with less unsuspected amount. The other second mechanism is to pay in cash for services that allow paper money. With the help financial workers to increase amount of money deposited to the financial institutions. Aborted transactions are funds hold by lawyer with some reason then this limit is lifted then returned to the client. The second stage layering is done after placing the money in banks and to avoid tracing even

the small transactions again divided to other smaller transaction. The final stage integration is receiving the money from the banks in legal way without attention of law and tax experts. If they caught by giving bribery to experts to reduce the tax payment. To back out the money they use fake workers. After that investing the money in luxury assets like ship, house and plane.

Machine Learning and Financial Sector

Machine learning is a technology that can be applied in different sectors including finance to modernize the sectors. It is the upcoming technology in developing countries can change the finance in different area, especially in fraud management. Machine learning has different uses in finance. In fraud management money launderings are the most important factor that affects financial sector to more insecure.

Machine learning techniques are used to increase efficiency and reduce fraud activities, also to solve various problems in the financial sector. Financial sector plays an important role in the good money exchange platform and world economy. With the rapid population growth and people's needs financial institutions for better financial systems, there is a constant need for this system to improve services and produce more profit. Machine learning method is about learning process or from previous experience. These methods need to be learnt to perform specific task. And have data that are based previous experience have set of attributes. These attributes defines the characteristics such as variable and features. The information on modern finance is based on detectors and devices that can better understand the environment like to improve security, financial monitoring, fraud management, account administration and financial system. It helps to identify fraud, monitor financial activities, robotic process automation that helps the sector in any financial activities from financial modeling up to invoice processing and wealth management. The financial institutions data will use to take quick and fast decision. In financial sector the most important machine learning mechanism is process automation using this can provide a better system by reducing operational cost and improving security.

5. Related Work

In this section, related works related to the study were reviewed and discussed. One of our objectives is to use the variety of data available of money laundering to develop anti-money laundering identifyion and identification model using machine learning techniques. In order to analyze this problem better, focus on the literature review.

Amr Ehab Muhammed Shokry et al. [9] conducted a study on money laundering detection using machine learning techniques. The study used a variety of libraries and tools like Numpy, Matplotlib, SKLearn and Pandas packages to carry out the investigations. The dataset was divided into two parts: training and testing in the ratio of 90:10. The researcher used one class SVM and Isolation forest for building detection model. In their study they approved that for dataset of 118,250 the Isolation forest algorithm is better than one class SVM. Conducting and injecting different suspicious activity are controlled and defined by domain expertise. Transactions behave like normal bank account transfer money to other accounts and a small fraction of accounts make suspicious transactions modeled on real world patterns. When the anomalies of the dataset increases the accuracy of the model decreases. The model used to identifying and detecting, hidden pattern and structuring of the grouping and classification. The Isolation Forest algorithm has been tested on transactions in different types got very successful detection accuracy.

Jos´e-de-Jesús Rocha-Salazar, María-Jesús Segovia-Vargas, et al. [10] designed a machine learning model using neural networks and abnormality indicator to detect and identify money laundering and terrorism in finance sector. The experiment was conducted using fuzzy logic for assigning risk metrics and C-means and neural gas clustering for transaction clustering neural network money laundering detection. Researchers used 30,278. Because the algorithm they had used for clustering can group a single transaction into different clusters, analyzing each cluster's transaction increases confusion between similar transactions. And selection of algorithm and testing different algorithms will make better accuracy to identify the highest accuracy.

The study entitled "Machine learning methods to detect money laundering in the Bitcoin blockchain in the presence of label scarcity" [11] has a scope of recognizing illicit patterns to increase the detection of money laundering in bitcoin transaction dataset. The study also identified that the insufficiency of data in the field which are labeled data. From the data of cryptocurrencies to train supervised classifier.

The blockchain of bitcoin includes 49 graphs that represents initial of one transaction to following related transactions with dataset of 203,769 transactions. To detect the illicit transaction, they used seven machine learning algorithms to design the model from supervised algorithm random forest is better performance with high F1-score. The active learning is outstanding as the imbalance of the data increases. The paper proposes a good approach to recognize and detect money laundering in blockchain of bitcoin in a complex unlabeled data scene with better fraud detector algorithms.

Yilma Goshime [12] built a detection model on financial fraud detection using machine learning algorithms (J48, ANN and Multilayer perception for data mining) with data mining concepts. The study used a dataset from 7500 financial datasets 5222 financial records. The tool used to design the model was MATLAB with better classification accuracy.

Abolfazl Mehbodniya, et al. [13] designed financial fraud detection for healthcare using machine and deep learning techniques. For identification process data are collected from different open source and combined together around 30000 customer's data. From these data s 70% is for train and 30% for test have been used. The algorithms where used to design a model are five; Random Forest has the highest accuracy (97.58%).

6. Research Methodology

Research design is a flow of task designed to answer research questions. So, in order to answer the research questions, the flow of the study have been designed as depicted in Figure 1. This study followed an experimental research approach, which included identifying research objectives and building machine learning models to validate the concept. Experimental research design approaches find the relationship between two variables i.e., independent variables and dependent variables in data samples [14].

In the area of machine learning data is crucial to answer every problem especially in data driven models. In this study, to build a model that can identify the fraud transaction will be attempted. First of all, the data for the building of this model will be collected from Commercial bank of Ethiopia. The Centre is found in Addis Ababa that located in leghar around. In [15] stated, the commercial bank of Ethiopia was established in 1942 by the name of state bank of Ethiopia, later then legally changed the name in 1963. Currently CBE has more than 23 million accounts with

more than 1100 branches. The study supports the fight against fraud especially in money laundering and the improvement of identification of illegal money transaction patterns.
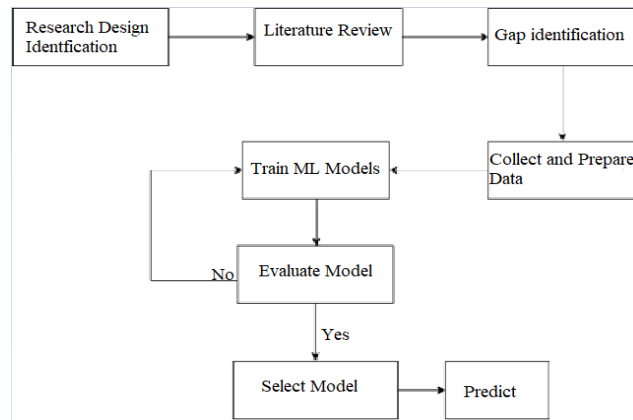

Figure 1. Research design of the study

The dataset used for this study contains 14740 records and each record have 13 features. From the total features there are seven categorical features (class), three are numerical values, and 3 Date type values. The dataset describes in the following tables.

6.1. Data Preprocessing

In the era of data analytics, machine learning and artificial intelligence, data is seen as the backbone. Real world data is "inconsistent". This means that the data is incomplete, noisy, inaccurate (including errors and outliers), and not of high quality, and as a result the model cannot produce quality mining results. Data preparation can affect a model's identify ability. Preprocessing the data can unexpectedly improve the accuracy of your model.

Missing values and noise are present in the AML dataset. This is due to data cleaning, categorical and numeric missing value handling, nominal to numeric conversion, feature selection, and other preprocessing tasks. For this study, we clean up the data set according to the following data preprocessing architecture:
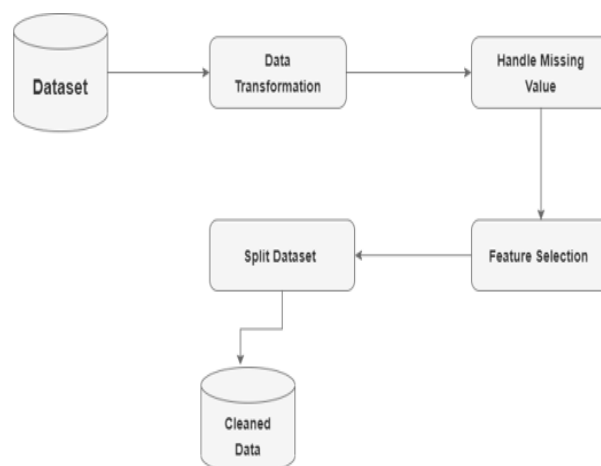

Figure 2. Data preprocessing architecture

These section mainly focused on missing data to prepare data for the model in these steps. By filling in missing data, we can clean data for a model. There are many options for preparing data based on data type and the amount of

https://iijsr.com

data missing. Such techniques include mean, median, most frequent value, and regression. These listed ways of filling may affect data quality by filling a similar deal for every missing value. In this study, a mean and mode is used to fill missing values.

Our dataset contains some attributes with Categorical data, some attributes with Numerical data machine learning models only manipulate with numbers. So we transform all of the Categorical data into numerical data. For example, Transaction Type Cell values: CASH DEPOSIT = 0; cash withdrawal= 1, withdrawal = 2, and Deposit= 3. Gender value: M=1, F=0.

Missing and/or corrupted data often contaminates bank transaction data. It occurs when there is no stored data value for the variable in an experiment or test. The absence of data is a common occurrence that can significantly affect the conclusions drawn from the data. Transaction data often contains many missing values, which makes analysis difficult for researchers who want to build a model using those data. transaction data have missing values for various reasons; when the customer does not fill in the all required information and the bank clerk does not enter all the correct information, it is often difficult to collect complete data.

The dataset is created from the data gathered from bank transaction data, which contains missing values and the missing data cannot be deleted because important information used in this study might be missed. And most features have missing values as shown Figure 2 because of this deleting the features that have missing values not allowed.

This study uses mean and mode imputation method which used to replacing the missing data with frequency and average values. Deleting any columns or rows that has any missing value may affect the training of machine, the above approach applied for false identifyion, by replacing the missing data with columns or rows value estimated by other similar data. This study will use Mode method which is used to impute or replace categorical features missing value and Mean to impute or replace numerical features.

6.2. Feature Scaling

These section mainly focused on missing data to prepare data for the model in these steps. By filling in missing data, we can clean data for a model. There are many options for preparing data based on data type and the amount of data missing. Such techniques include mean, median, most frequent value, and regression. These listed ways of filling may affect data quality by filling a similar deal for every missing value. In this study, a mean and mode is used to fill missing values.

Our dataset contains some attributes with Categorical data, some attributes with Numerical data machine learning models only manipulate with numbers. So we transform all of the Categorical data into numerical data. For example, Transaction Type Cell values: CASH DEPOSIT = 0; cash withdrawal= 1, withdrawal = 2, and Deposit= 3. Gender value: M=1, F=0.

Missing and/or corrupted data often contaminates bank transaction data. It occurs when there is no stored data value for the variable in an experiment or test. The absence of data is a common occurrence that can significantly affect the conclusions drawn from the data. Transaction data often contains many missing values, which makes analysis

difficult for researchers who want to build a model using those data. transaction data have missing values for various reasons; when the customer does not fill in the all required information and the bank clerk does not enter all the correct information, it is often difficult to collect complete data.

The dataset is created from the data gathered from bank transaction data, which contains missing values and the missing data cannot be deleted because important information used in this study might be missed. And most features have missing values as shown Figure 2 because of this deleting the features that have missing values not allowed.

This study uses mean and mode imputation method which used to replacing the missing data with frequency and average values. Deleting any columns or rows that has any missing value may affect the training of machine, the above approach applied for false identifyion, by replacing the missing data with columns or rows value estimated by other similar data. This study will use Mode method which is used to impute or replace categorical features missing value and Mean to impute or replace numerical features.

6.3. Feature Selection

Feature selection is a way of selecting relevant attributes from given dataset attributes[16].At this stage the study will be identifying the related features from a set of data and removing the irrelevant or less important features with do not contribute much to the target variable in order toachieve better accuracy for the model. Several Researchers uses different types of feature selection method these include Filter method, Wrapper method and Embedded Method.

One of the measures used for feature selection is dependency measures. Many dependencies based methods have been proposed. The main measure is Correlation based method. Pearson's Correlation method is used for finding the association between the independent features and the dependent feature. This study uses correlation method for feature selection. Correlation Method evaluates all attributes related to the target class via Pearson's correlation method, which provides a ranking of the attributes from high to low and this reduces processing time and the data dimension [17].

6.4. Evaluation

To evaluate the performance of the identifyion model; by using accuracy, specificity, sensitivity, and confusion matrix: precision, recall, F1. Calculating true positive (TP), True Negative (TN), false Positive (FP), and false Negative (FN).

7. Result

The total number of records in the collected dataset, after preprocessing, was still 14740. As a result, the dataset cannot lose any data bank transactions, and the simpleimputation() technique is used to properly handle the dataset's missing values. 3355 transactions out of the total dataset had fraud, while the rest 11385 have normal transactions. The imbalance dataset may varies the result of identifyion the selected algorithm classifier parameters. It lead the identifyion to incorrect and the training of the machine only be the major classes. The oversampling method was employed to strike a balance between the minority class's and the majority class'

respective values. Following the application of the resembling approach, the amount of the class labels for the classification class attributes was balanced to 9130 normal and 9130 fraud transactions, totaling 18,260. In this study, the approach of Pearson correlation analysis was used to choose the features.

Due to the correlation analysis function's defined threshold value, which in this case is greater than 70% correlated one feature from the other feature, it is not possible to remove any features from the dataset using this method as a result because it assumes that both features have the same values, which would prevent removing one of the two features. However, there is not enough feature correlation in this case to satisfy the threshold parameter. There is a 39% association between balances kept and account numbers, and there is a 23% correlation between account type and occupation. Negative correlation is taken into account in this study when defining the absolute function because it is also crucial and indicates when a feature's value is increasing while another value is decreasing on the opposite side.

8. Conclusion and Recommendation

8.1. Conclusion

Money Laundering is a huge problem that affects individuals, especially in governmental stage and disturbance of day-to-day activities. It is a major challenge in developing countries like Ethiopia, to manage and control these kinds of problem it needs more time and resource that makes the problem more difficult and this identifyion model helps the fraud investigators and the bank management system to easily identify the fraud problem and the progress of the fraud problem by maintaining the pattern of the money laundering. This study attempt to build a model for Anti-money laundering identifyion and identification using machine learning techniques. The first step in machine learning is collecting data. Collecting data is not an easy task because it may have missing value, unnecessary data and inappropriate to use it to train the machine directly. For data preparation used to train machine learning models and the techniques used for preparation called data preprocessing. Data preprocessing is used to clean, handle missing value, encoding data for the model to be trained. In data preprocessing the main procedure is to convert the collected data to machine understandable methods. The features in dataset having above 50% missing value are eliminated not filled by other handling missing value techniques. In this study the features with missing values was handled using mean/mode imputation, for the numerical variables the mean imputation used and the mode imputation for categorical variables.

In any machine learning studies, there will be imbalanced datasets. In this study also, the dataset have imbalanced class distribution. Training the model by this imbalance dataset may led the identifyion result of the model to be incorrect and Synthetic Minority Oversampling Technique analysis used to prevent the wrong result occurred in the identifyion. The model is developed using this clean dataset. After using Synthetic Minority Oversampling Technique to the dataset, the cleaned data was 18260 and balanced 9130 normal and 9130 fraud transaction. The feature selection applied on preprocessed dataset was Pearson Correlation analysis. For this study five machine learning algorithms was selected to develop the identification model and evaluate it by confusion matrix. Using the models on the cleaned dataset, the accuracy results of this study from the selected algorithms respectively, decision tree was 92.9%, support vector machine was 91.5%, logistic regression was 89.8%, Naïve Bayes 82.8% and

random forest get accuracy 99.1%. Relatively from each model the Random Forest model for gives a better identifyion.

8.2. Recommendation and Future work

Some future work identified throughout the research may be carried out. Here in this study only one district of the Commercial Bank of Ethiopia dataset was explored and analyzed, in the future, another district of the Commercial Bank of Ethiopia dataset can be explored, further research is needed to handle additional variables. The model is also recommended to train and test on a large number of data with complex configurations. The five machine learning techniques were used in this study on the Commercial Bank of Ethiopia dataset; further other techniques can be explored as well; different machine learning algorithms can be explored, and data can be analyzed.

This study contribution could be used to design and develop a identification model to identify illegal transaction. Finally, based on the study's findings, the study has the following recommendations.

• This study used different classification algorithms in which Random Forest Algorithm better than others Naive Bayes, logistic regression, Decision Tree and Support Vector Machine, so this study recommends the Random Forest model to build the Anti-money laundering identification model.

• The findings of this study can be used as input by the company, which can be integrated with its attributes to improve service delivery while reducing the risk of illegal transaction.

• Since the data preparation stage takes a long time, it is recommended that large companies, such as Banks, implement or own a data warehouse where all of their data can be formally stored based on basic features and where different machine learning techniques can be used to simplify day-to-day service provisioning.

References

[1] S. Report (2021). En 2021 13.

[2] Machine Learning in Anti-Money Laundering-Summary Report, Pages 1–10, 2018.

[3] OECD (2019). Money Laundering and Terrorist Financing Awareness Handbook for Tax Examiners and Tax Auditors, OECD, Paris.

[4] M.Alkhalili, M.H.Qutqut, & F.Almasalha (2021). Investigation of Applying Machine Learning for Watch-List Filtering in Anti-Money Laundering. IEEE Access, 9: 18481–18496.

[5] J.Chen & S.Anderson (2022). Money laundering. www.investopedia.com/terms/m/moneylaunde ring.asp.

[6] Financial Accountability Transparency & Integrity Panel, "FACTI Panel Interim Report," September 2020.

[7] B. Herron and S. Duggal (2020). Deploying Machine Learning Models in an Anti-Money Laundering (AML) Program. SAS Glob. Forum, Pages 4553–2020.

[8] A.J. Kidwai (2006). Money Laundering and the Role of Banks. Pakistan Horiz., 59(2): 43–47.

[9] A.E.M. Shokry, M.A. Rizka, and N.M. Labib (2020). Counter terrorism finance by detecting money laundering hidden networks using unsupervised machine learning algorithm. Proc. 13th IADIS Int. Conf. ICT, Soc. Hum. Beings 2020, ICT 2020 Proc. 6th IADIS Int. Conf. Connect. Smart Cities 2020, CSC 2020 Proc. 17th IADIS Int. Conf. , Pages 89–97. doi: 10.33965/ict_csc_wbc_2020_202008l012.

[10] J. de J. Rocha-Salazar, M.J. Segovia-Vargas, and M. del M. Camacho-Miñano (2021). Money laundering and terrorism financing detection using neural networks and an abnormality indicator. Expert Syst. Appl., 169. doi: 10. 1016/j.eswa.2020.114470.

[11] J. Lorenz, M.I. Silva, D. Aparício, J.T. Ascensão, and P. Bizarro (2020). Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity. ICAIF 2020 - 1st ACM Int. Conf. AI Financ. doi: 10.1145/3383455.3422549.

[12] Yilma Goshime (2017). Constructing financial fraud detection model using machine learning technique: the case of Wedera farmers multipurpose cooperative union, 549: 40–42.

[13] A. Mehbodniya et al. (2021). Financial Fraud Detection in Healthcare Using Machine Learning and Deep Learning Techniques. Secur. Commun. Networks., doi: 10.1155/2021/9293877.

[14] S. Bell (2009). Experimental Design. Int. Encycl. Hum. Geogr., Pages 672–675, doi: 10.1016/B978-0080 44910-4.00431-4.

[15] CBE (2020). Commercial Bank of Ethiopia Environmental and Social Policy.

[16] M. Shardlow (2016). An analysis of feature selection techniques. Univ. Manchester, 1: 1–7.

[17] E. Bahadır (2016). Using neural network and logistic regression analysis to identify prospective mathematics teachers' academic success upon entering graduate education. Kuram ve Uygulamada Egit. Bilim., 16(3): 943–964. doi: 10.12738/estp.2016.3.0214.