International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Fraud Detection: Leveraging Deep Learning Methods for Retail Transactions

Mr. Kiran Maka *[1], Dr. S. Pazhanirajan [2], Dr. Sujata Mallapur [3]

Submitted: 19/07/2023 **Revised**: 10/09/2023 **Accepted**: 23/09/2023

Abstract*:* In this work, deep learning models are used to expect the swindle in monetary declarations and to derive the important features that will be used by the auditors to control the deception in the reported declarations. Totally, eight models are developed as part of this work. For under sampled and oversampled datasets, the models are built on deep neural networks and convolutional neural networks, each having three and five layers. Top two models from the eight models are selected based on performance factors like accuracy, sensitivity and precision. The precision is defined for both positive predictive value and negative predictive price. From the top two models, important features are derived using the SHAP methodology. The important features from both the top models are analyzed to derive a common set of features that would be recommended to auditors to make their job easy and accurate.

Index Terms: *Economic statements, fraud, machine learning, data mining*

1. Introduction

The financial statements prepared by accounting and finance departments are issue to scrutiny by the controlling bodies by SEBI and RBI in order to ensure the genuineness as the economy of a nation depends on financial health of the companies and safety of investors wealth. Financial statements reflect the financial condition, investment details, liabilities, interest paid and interest earned etc. The financial condition [1] of an organization reveals correct information about the growth of the assets or liabilities over time. The financial statements are used by the organizations for the purpose of raising new loans or to seek investments in the markets. Also, these statements are used by investors to decide on their investments, by rating agencies for credit ratings, by creditors to grant or recover loans, or by the governments or third parties to recognize the good performance with awards.

Since the financial statements are important documents used by the external entities to gauge the financial health of an organization, there is possibility of manipulation of the certain figures in the financial statements by the companies to make it more attractive for new investments, loans or for awards. Of late, there is growing trend [2-4] of fraud in the emerging markets. Therefore, it is imperative to ensure the genuineness of the statements before using it for taking decisions. The genuineness can be verified using machine learning models. These machine learning models are mostly cataloguing models which are trained on a set of labeled financial statements. The labels of Genuine or Fraudulent are provided to each financial statement by carefully studying by the expert auditors. When an organization employs digital transformations, these models are made part of the transformation and it becomes an important tool to classify the financial statements.

Models are built not just for determining the presence of fraud in the statements but also to derive the explanability about why these statements are classified as fraudulent. The important factors driving the fraud must be understood thoroughly so that these factors can be used by the auditors when conducting the manual audits. While the procedure to build a model for the purpose of classification for fraud or for deriving important factors is same, an extra algorithm like SHAP is run over the model to derive the important factors. There are many research works happened in the past to identify the best algorithm to check for the fraud in the statements [5-6].

Many a times, the algorithms used for classification purpose of any problem are generally logistic regression, neural networks, xgboost or adaboost models. These models chosen for the required accuracy of classification and it also depends on the dataset. However, not all algorithms are explainable. Some models like neural networks, xgboost or adaboost are black box models. It only provides the probability of classification with good accuracy [7], but it does not explain which factors are driving the classification label for a given record. Hence, these models are termed as black box models. There are

[1]*Research Scholar, Department of CSE, Annamalai University, kiranmaka@gmail.com*
[2]*Assistant Professor, Department of CSE, Annamalai University, pazhanisambandam@gmail.com*
[3]*Professor, Department of CSE, Sharnbasva University, Kalaburagi Karnataka, sujatamallapur@gmail.com*

additional approaches like support vectors [8] and Zipf's law [9] that provide classification with good accuracy. Some of the previous works that happened in the past are based on neural networks and belief networks [4,10], neural networks using back propagation algorithm [11,12,15,16], risk assessment based neural networks [13], methods to preprocess the data in preparing it for training the fraud detection models [14], importance of distribution of digits [17] and other important methods [18-20].

There is significant analysis done in arriving at the right model and important factors by the authors [21,22]. Authors tried 38 models for different combinations of datasets. The algorithms like regression and tree-based methods were used. Authors established a methodology to originate the significant structures driving the uncovering of fraud. The dataset was split and enhanced into two variants, namely, over sampled dataset and under sampled dataset. The 38 models were trained with a distinct algorithm and an oversampled or under sampled dataset for each model.

However, most of the models built by the researchers in the past were based on machine learning. The features were derived manually and was input to classification layer to determine the probability. The accuracy of prediction depends on type of features present and feature engineering in a machine learning model. The dependency on feature engineering can be overcome by employing deep learning models and in the contemporary work, an effort is made to detect the fraud and derive important features using deep learning algorithms like deep neural networks (DNN) and convolutional neural networks (CNN). The DNNs were developed overtime by many researchers [23-27] and CNNs were introduced in [28]. There are many variants of CNNs which are popular in the space of image processing which can be applied to other fields like fraud detection as well as natural language processing [28-40].

In Sec. II, modeling methodology and steps followed in the preparation of dataset are explained in detail. In Sec. III, simulation results are provided and deriving of important features are explained. In Sec. IV, important points are explained and results are discussed along with conclusions.

2. Dataset and ML models

In this unit deep learning approaches are discussed along with the data used in training the models. Deep learning model require big data as contrast to machine learning, to automatically extract the features from the data. In machine learning the features required to classify a given record or data are extracted manually by applying various feature extraction techniques. In deep learning methods, the architecture of a neural network has two parts. The first part of the architecture extracts the features and the second part of the architecture classify the processed data. In other words, the architecture may be split into feature extraction layer and classification layer. In case of DNN, both feature extraction layer and classification layers are constructed with fully connected layers. In case of CNN and other CNN based models, only the classification layer is constructed with fully connected layers. The feature extraction layers of CNN based models are usually constructed out of convolution and pooling layers.

In order to shape a model, the architecture of the model and the data are important. The data represents the information and patterns to be extracted using the model. Model must be robust and effective enough to extract the information efficiently.

The data used in this work has been extracted from 14000 statements or financial reports. The 14000 statements are reported by 3500 firms in a time period of 5 years. There are multiple reports by some firms as firms report the statement almost every year. For the purpose of feeding the financial statement to the computer, the data present in the financial statements are mapped to tabular forms. Each row in the table characterizes a declaration of a secure in a year. When the data has been mapped to tabular form, the data does not take the information around if the financial declaration is unaffected or fake. A group of auditors have analyzed the financial statements and labelled each record as genuine or fraudulent [21]. Some of the statements which were marked by auditors as the labeling also indicates deviation, lack of adherence to principles, negative impact, or report as fake. It has been found that 358 out of 14000 statements are either fraudulent or near fraudulent cases. Remaining statements are treated as genuine in the table. With this, it can be concluded that data is highly imbalanced for the label as fraudulent records in the table are just 358 out of 14000. The items or components of the financial statement are entered as columns in the table. Each column is a variable in the data. Thus, there were 17 variables in the table. Important variables included

Interest earned (interest_earned)

Altman Z-Score (az_score)

Ratio of Total debts to Total assets (tdebts_tasset or td_ta)

Ratio of Debt to Equity (debt_equity)

Total Assets (tasset or ta)

Total Liability (tliability or tl)

Return on Equity (roequity or roe)

Ratio of Total Accruals to Total assets (total_accruals_ta)

Ratio of Investment to Sales (inv_to_sales)

Ratio of Total sales to Total assets (sales_tassets or sales_ta)

Ratio of accounts receivables to sales (ac_recv_to_Sales)

Beniesh M-Score (m_score)

Total Sales (sales)

Total accounts receivables (ac_recvbl)

Ratio of PPE and Total assets (ppe_tasset)

Ratio of Fixed asset and Total asset (fixedAsset_tasset)

Gross Margin (gross_margin).

As described earlier, there are only 358 fraudulent statements in a dataset of 14000 records. This represents a skewed dataset or imbalanced dataset for the label. The skewness in the label distribution can be corrected to have 50% for each of the classes. The dataset can be modified to have approximately 50% of the genuine class and remaining as fraudulent class. The dataset can be corrected to have a balanced labels using two methods, namely, under sampling and over sampling. In case of under sampling, the records of genuine class are sampled randomly. The amount of annals sampled out of the genuine class is equal to the number of minority class. In this case the minority class is the fraudulent class. Therefore, 358 records are selected randomly from the population of 13,642 records of genuine class. With this approach, though the classes are balanced, the information and patterns available in the remaining records of the majority class are never used. In oversampling approach, the minority class is experimented manifold aeras repeatedly to match the count of records in popular class. In this approach, the 358 records of fraudulent class replicated multiple times randomly to match the count of 13,642 records of genuine class. With this approach, the information and patterns present in the minority class is duplicated multiple times which may lead to oversupply of patterns to the model. Both the under sampling and over sampling have their own advantages and limitations. In this work, both the under sampled and oversampled datasets are used to build the models to derive the important features driving the prediction.

However, not all of the 14000 records of the data are of good quality in terms of completeness and hence the only those records that were of good quality were selected for training and testing purpose. A set of 4,960 records were selected for training and 1,240 records for following cleaning, testing. By applying the oversampling procedure, the size of oversampled dataset became 9656 records and that of under sampled dataset became 264 records. Each of these datasets has been balanced to have 50% class distribution.

Several models were trained using both under sampled (US) and over sampled (OS) data. The models that were trained in this work are based on neural networks or deep learning-based models.

- N1 Neural Network Model – 3L - DNN – US
- N2 Neural Network Model – 5L - DNN – US
- N3 Neural Network Model – 3L - CNN – US
- N4 Neural Network Model – 5L - CNN – US
- N5 Neural Network Model – 3L - DNN – OS
- N6 Neural Network Model – 5L - DNN – OS
- N7 Neural Network Model – 3L - CNN – OS
- N8 Neural Network Model – 5L - CNN – OS

From these eight models, two best models are selected based on the performance criteria. In the present work, accuracy of models, precision and sensitivity are selected as performance criteria. The importance features driving the prediction are extracted from the model using game theory based SHAP method.

Algorithm proposed:

1. Manually label the dataset and generate M-scores.
2. Extract train and test datasets
3. Create oversampled train set and under sampled train set.
4. Train the eight neural network based deep learning models with oversampled train set and under sampled train set
5. From the eight models, select two best models based on the performance criteria of accuracy, sensitivity and precision. Of the two models, one model is to be chosen based on under sampled train set and the other on over sampled train set.
6. From these two top models, extract the important features using SHAP.
7. Select the top 5 significant features from each of these two models.
8. Select common features appearing in both these models.
9. Use the common features as a guideline to check for fraudulency in the statements.

3. The Simulation results

In this segment, performance results of the eight models are presented and discussed. For the eight models, performance metrics like precision, sensitivity, specificity, positive predictive value, and negative predictive value are computed. The results are compared with each other and best two models out of eight models are selected.

Table 1: Performance metrics of deep learning models

Model Index	N1	N2	N3	N4	N5	N6	N7	N8
Model Name	3L - DNN – US	5L - DNN – US	3L - CNN – US	5L - DNN – US	3L - DNN – OS	5L - DNN – OS	3L - CNN – OS	5L - DNN – OS
	Undersampled Dataset				Oversampled Dataset			
Dataset	Training	Training	Training	Training	Training	Training	Training	Training
TP	73	72	75	74	2693	2564	2623	2396
FN	20	21	18	19	682	811	752	979
TN	68	64	70	69	2404	2301	2234	2507
FP	24	28	22	23	971	1074	1141	868
Accuracy	0.762	0.735	0.784	0.773	0.755	0.721	0.720	0.726
Sensitivity	0.785	0.774	0.806	0.796	0.798	0.760	0.777	0.710
Specificity	0.739	0.696	0.761	0.750	0.712	0.682	0.662	0.743
Pos Pred Value	0.753	0.720	0.773	0.763	0.735	0.705	0.697	0.734
Neg Pred Value	0.773	0.753	0.795	0.784	0.779	0.739	0.748	0.719
Dataset	Testing	Testing	Testing	Testing	Testing	Testing	Testing	Testing
TP	31	29	30	29	1097	1033	1054	1098
FN	9	11	10	11	356	420	399	355
TN	29	27	33	32	1032	1211	1034	1026
FP	10	12	6	7	421	242	419	427
Accuracy	0.759	0.709	0.797	0.772	0.733	0.772	0.719	0.731
Sensitivity	0.775	0.725	0.750	0.725	0.755	0.711	0.725	0.756
Specificity	0.744	0.692	0.846	0.821	0.710	0.833	0.712	0.706
Pos Pred Value	0.756	0.707	0.833	0.806	0.723	0.810	0.716	0.720
Neg Pred Value	0.763	0.711	0.767	0.744	0.744	0.742	0.722	0.743

From the above table, it can be observed that model N3 out performed on training data with 78.4% accuracy among the under sampled databased models as shown in Fig. 1. Similarly, for the over sampled data, model N5 is the best with 75.5% among the four models. When models were tested on the testing data, again model N3 has outperformed the under sampled based models with 79.7% as shown in Fig. 2. With over sampled test set, the model N6 outperformed the remaining over sampled based models with 77.2% accuracy.

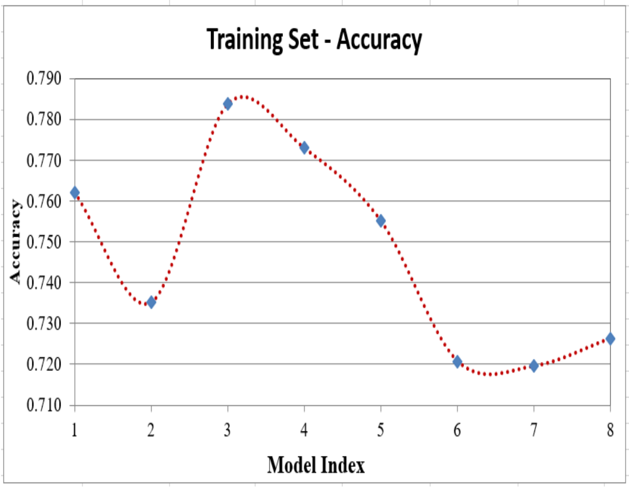


Fig 1: Accuracy of Neural network – deep learning models on training set

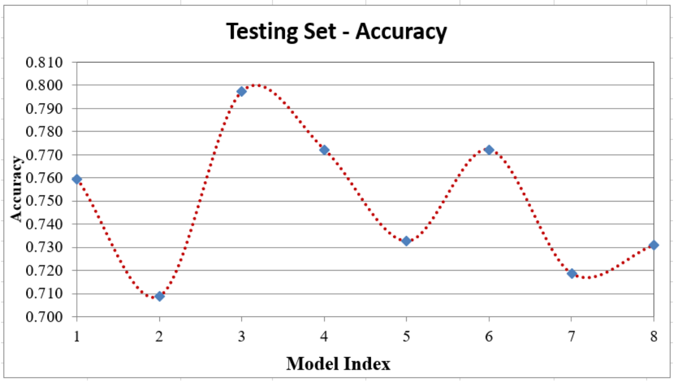
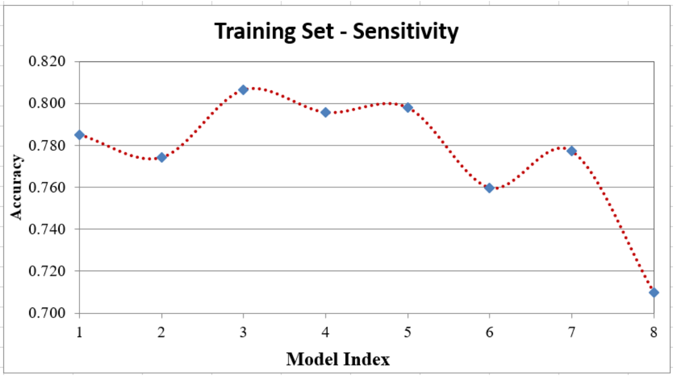


Fig 2: Accuracy of Neural network – deep learning models on testing set

When sensitivity is compared, again model N3 has scored 80.6% on the under sampled training set and model N5 has scored 79.8% among the over sampled training set as shown in Fig. 3. On the testing set, as shown in Fig. 4, model N1 is the best with 77.5% for under sampled data and model N8 is the best with 75.6% for over sampled data.



Figire 3: Sensitivity of Neural network – deep learning models on training set

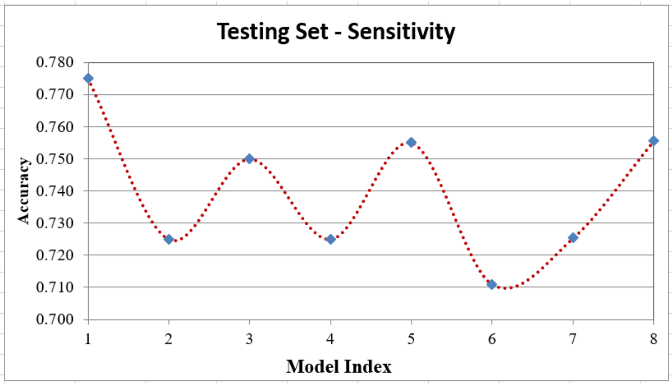


Fig 4: Sensitivity of Neural network – deep learning models on testing set

When specificity is compared, model N3 has scored 76.1% on the under sampled training set and model N8 has scored 74.3% among the over sampled training set. On the testing set, model N3 is the best with 84.3% for under sampled data and model N6 is the best with 83.3% for over sampled data.

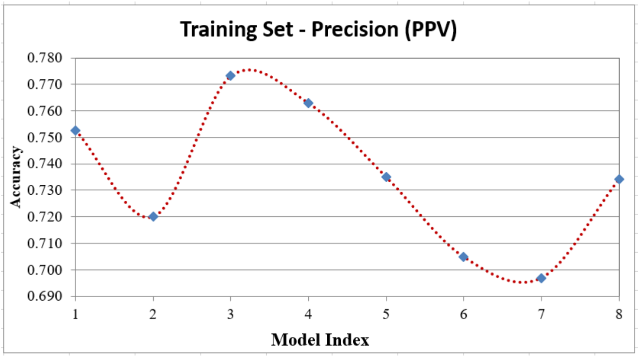


Fig 5: Accuracy of Neural network – deep learning models on training set

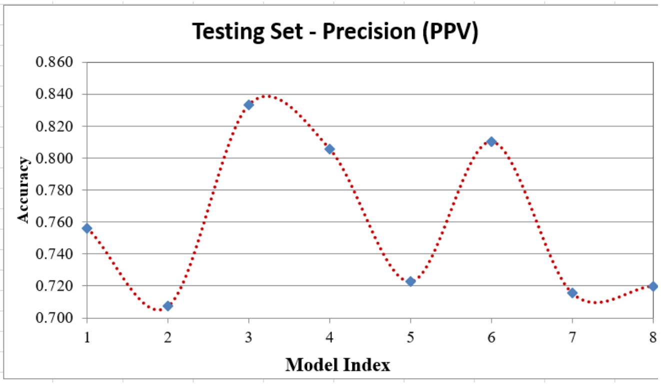


Fig 5: Accuracy of Neural network – deep learning models on testing set

For the precision or positive predictive value, model N3 out performed on training data with 77.3% accuracy among the under sampled databased replicas as revealed in Fig. 5. Similarly, for the over sampled data, model N5 is the best with 73.5% among the four models. When models were tested on the testing data, again model N3 has outperformed the under sampled based models with 83.3% as shown in Fig. 6. With over sampled test set, the model N6 outperformed the remaining over sampled based models with 81.0% accuracy. For negative predictive value, models N3 and N5 are good with 79.5% and 77.9% for under sampled and over sampled training sets respectively. Similarly, models N3 and N6 are good with 76.7% and 74.4% for under sampled and over sampled test sets respectively.

From the above analysis, model N3 has outperformed both on training and testing under sampled data sets for most of the metrics. In case of over sampled dataset, the model N5 is good for training and model N6 is good for testing data, Since the model performance of the testing data is preferred over the training data, model n6 is selected as the best among all the models on the oversampled dataset.

The model N3 corresponds to 3L-CNN-US and model N6 corresponds to 5L-DNN-OS. It is quite clear that the CNN model with three layers has performed well on the under sampled data and DNN model with five layers has done well on oversampled data.

Table 2: Imperative characteristics derived by 3L-CNN-US model

Feature status rank – N3	Feature term – N3	Importance score – N3
1	interest_earned	100
2	az_score	92.11
3	tdebts_tasset	67.646
4	tasset	48.814
5	debt_equity	48.229
6	tliability	34.823
7	total_accruals_ta	33.759
8	roequity	29.46
9	inv_to_sales	28.636
10	sales_taseets	26.455
11	ac_recv_to_Sales	22.18
12	m_score	21.532
13	sales	16.85
14	ac_recvbl	13.023
15	ppe_tasset	11.537
16	fixedAsset_tasset	8.617
17	gross_margin	0

Table 2 shows the important features derived from the 3L-CNN-US model using SHAP. The importance score of the model is also shown in the table. It can be observed from the table 2 that interest_earned, az_score, tdebts_tasset, tasset, debt_equity and tliability are the top six features driving the forecast for spotting fraud in financial statements. The importance score of the variables provided in the table do not carry the weightage of the variable in determining the fraud, rather it is only an indicative number to measure the position of the variable with respect to other variables.

Table 3: Imperative characteristics derived by 5L-DNN-OS

Feature importance rank – N6	Feature name – N6	Importance score – N6
1	interest_earned	100
2	az_score	65.79
3	tdebts_tasset	56.77
4	tliability	45.63
5	debt_equity	60.66
6	tasset	28.32
7	ac_recv_to_Sales	21.44
8	inv_to_sales	21.23
9	gross_margin	19.65
10	fixedAsset_tasset	19.33
11	total_accruals_ta	18.45
12	sales	17.45
13	sales_tasset	17.23
14	roequity	15.63
15	m_score	14.25
16	ac_recvbl	13.41
17	ppe_tasset	0

Table 3 shows the list of important features along with its score. These scores were derived from the SHAP methodology and it carries the value of relative positions in terms of significance in predicting the fraud in the statements. Variable quantity like interest_earned, az_score, tdebts_tasset, tliability, debt_equity and tasset are the top six features driving the prediction to perceive the deception in the economic declarations

Table 4: Top 6 Imperative characteristics derived from deep learning models

Feature importance rank	Feature name
1	interest_earned
2	az_score
3	tdebts_tasset
4	tliability
5	debt_equity
6	tasset

By comparing the top six variable that play important role in determining the deception in the economic reports, the variables interest_earned, az_score, tdebts_tasset, tasset, debt_equity and tliability are common between the two models N3 and N6 though the order is different. The order of the top three variable is same between N3 and N6 and position of debt_equity is also same. The only difference between the N3 and N6 in terms of order is tliability and tasset. In N3 model, the tliability is at position 4 and tasset is at position 6. In case of N6 model, it is just reverse, that is tliability is at position 6 and tasset is at position 4. Hence the top six variables are reorganized as aligned to N6 model since the model is built with oversampled data and number of data points are much higher compared to an under sampled dataset. Therefore, the top 6 variables, shown in Table 4 are selected to provide a supervision to the examiner in deciding the presence of con in the statements.

4. Conclusion

In this study, two models are chosen from a pool of eight to determine the key characteristics that influence the forecasting of financial statement fraud. All the eight models were built with deep learning models. The models were built with combinations of DNN, CNN, 3 Layers, 5 Layers, Under sampled and over sampled data sets. Two models were selected from each under sampled and over sampled, 3 layer and 5 layers, DNN and CNN. The two models are 3L-CNN-US and 5L-DNN-OS. These two models were selected from the eight models by analyzing accuracy, sensitivity, specificity, positive predictive value and negative predictive values. The two models were run with SHAP algorithm to derive Imperative characteristics that drive the predictions. It has been found that interest_earned, az_score, tdebts_tasset, tasset, debt_equity and tliability are the top six features driving the prediction to detect the fraud in the financial statements using 3L-CNN-US. The values for interest_earned, az_score, tdebts_tasset, tliability, debt_equity, and tasset are same are the top six features driving the prediction to perceive the con in the economic reports using 5L-DNN-OS. After comparison between both the models and analyzing, top six variables are reorganized as aligned to N6 model since the model is built with oversampled data and number of data points are much higher compared to an under sampled dataset.

References

[1] W.H. Beaver, Financial ratios as predictors of failure, Journal of Accounting Research 4 (1966) 71–111.

[2] http://en.wikipedia.org/wiki/Enron.

[3] http://en.wikipedia.org/wiki/MCI_Inc.

[4] Kirkos E, Spathis C, Manolopoulos Y. Data Mining Techniques for the Detection of Fraudulent Financial Statements. *Expert Systems with Applications* 2007; **32**: 995–1003.

[5] Glancy FH, Yadav SB. A Computational Model for Financial Reporting Fraud Detection. *Decision Support Systems* 2011; **50**: 595–601.

[6] Jans M, Werf JM, Lybaert N, Vanhoof K. A Business Process Mining Application for Internal Transaction Fraud Mitigation. *Expert Systems with Applications* 2011; **38**: 13351–13359.

[7] Neuroshell 2.0, Ward Systems Inc. http://www.wardsystems.com.

[8] Cecchini M, Aytug H, Koehler G, Pathak P. Detecting Management Fraud in Public Companies. *Management Science* 2010; **56**: 1146-1160.

[9] S.-M. Huang, D.C. Yen, L.-W. Yang, J.-S. Hua, An investigation of Zipf's Law for fraud detection, Decision Support Systems 46 (1) (2008) 70–83.

[10] Hoogs B, Kiehl T, Lacomb C, Senturk, D. A Genetic Algorithm Approach to Detecting Temporal Patterns Indicative of Financial Statement Fraud. *Intelligent Systems in Accounting, Finance and Management* 2007; **15**: 41-56.

[11] J.E. Sohl, A.R. Venkatachalam, A neural network approach to forecasting model selection, Information & Management 29 (6) (1995) 297–303.

[12] M.J. Cerullo, V. Cerullo, Using neural networks to predict financial reporting fraud: Part 1, Computer Fraud & Security 5 (1999) 14–17.

[13] T.G. Calderon, J.J. Cheh, A roadmap for future neural networks research in auditing and risk assessment, International Journal of Accounting Information Systems 3 (4) (2002) 203–236.

[14] E. Koskivaara, Different pre-processing models for financial accounts when using neural networks for auditing, Proceedings of the 8th European Conference on Information Systems, vol. 1, 2000, pp. 326–3328, Vienna, Austria.

[15] E. Koskivaara, Artificial neural networks in auditing: state of the art, The ICFAI Journal of Audit Practice 1 (4) (2004) 12–33.

[16] B. Busta, R. Weinberg, Using Benford's law and neural networks as a review procedure, Managerial Auditing Journal 13 (6) (1998) 356–366.

[17] E.H. Feroz, T.M. Kwon, V. Pastena, K.J. Park, The efficacy of red flags in predicting the SEC's targets: an artificial neural networks approach, International Journal of Intelligent Systems in Accounting, Finance, and Management 9 (3) (2000) 145–157.

[18] Lokanan, M.E. (2017), “Theorizing Financial Crimes as Moral Actions”, European Accounting Review, Vol. 0 No. 0, pp. 1–38.

[19] Lokanan, M.E. and Sharma, S. (2018), “A Fraud Triangle Analysis of the Libor Fraud”, Journal of Forensic and Investigative Accounting, Vol. 10 No. 2, pp. 187–212.

[20] Li, Z. (2016). "Anomaly detection and predictive analytics for financial risk management." Available at: https://rucore.libraries.rutgers.edu/rutgers-lib/49363/ (accessed 21 September 2018).

[21] Kiran Maka, S. Pazhanirajan, Sujata Mallapur, "Establishing Causality to Detect Fraud in Financial

[22] Statements", International Journal of Circuits, Systems and Signal Processing, Volume 15, 2021, DOI: 10.46300/9106.2021.15.166.

[23] K. Maka, S. Pazhanirajan and S. Mallapur, Selection of most significant variables to detect fraud in financial statements, MaterialsToday: Proceedings, https://doi.org/10.1016/j.matpr.2020.09.613

[24] W. S. McCulloch and P. Walter, "A logical calculus of the ideas immanent in nervous activity", in The bulletin of mathematical biophysics, vol. 5, no. 4, pp. 115-133, 1943.

[25] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain", in Psychological review, vol. 65, no. 6 pp. 386, 1958.

[26] M. Minsky and P. Seymour, "Perceptrons." 1969.

[27] D. H. Ackley, E. H. Geoffrey and J. S. Terrence, "A learning algorithm for Boltzmann machines", in Cognitive science, vol. 9, no. 1, pp. 147-169, 1985.

[28] K. Fukushima, "Neocognitron: A hierarchical neural network capable of visual pattern recognition", in Neural networks, vol. 1, no. 2, pp. 119-130, 1988.

[29] Y. LeCun, et al., "Gradient-based learning applied to document recognition" in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[30] G. E. Hinton, O. Simon and T. Yee-Whye, "A fast learning algorithm for deep belief nets", in Neural computation, vol. 18, no. 7, pp. 1527-1554, 2006.

[31] NAIR AND G. HINTON, “RECTIFIED LINEAR UNITS IMPROVE RESTRICTED BOLTZMANN MACHINES”, IN PROCEEDINGS OF THE 27TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML-10), 2010.

[32] G. E. Hinton and R. S. Ruslan, "Reducing the dimensionality of data with neural networks" in science, vol. 313, no. 5786, pp. 504-507, 2006.

[33] Krizhevsky, I. Sutskever and G. E. Hinton, “ImageNet classification with deep convolutional neural networks”, In NIPS, pp. 1106–1114, 2012.

[34] K. Simonyan and Z. Andrew, "Deep convolutional networks for large-scale image recognition", in arXiv preprint arXiv:1409.1556, 2014.

[35] M. Lin, C. Qiang and Y. Shuicheng, "Network in network" in arXiv preprint arXiv:1312.4400, 2013.

[36] J. T. pringenberg, et al., "Striving for simplicity: The all convolutional net" in arXiv preprint arXiv:1412.6806, 2014.

[37] Szegedy, et al., "Going deeper with convolutions", in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.

[38] Szegedy, I. Sergey and V. Vincent, "Inception-v4, inception-resnet and the impact of residual connections on learning" in arXiv preprint arXiv:1602.07261, 2016.

[39] K. He, et al., "Deep residual learning for image recognition" in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[40] G. Huang, et al., "Densely connected convolutional networks" in arXiv preprint arXiv:1608.06993, 2016.

[41] G. Larsson, M. Michael and S. Gregory, "FractalNet: Ultra-Deep Neural Networks without Residuals" in arXiv preprint arXiv:1605.07648, 2016.

[42] Mujawar, S. S. ., & Bhaladhare, P. R. . (2023). Effective Feature Selection Methods for User Sentiment Analysis using Machine Learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 37–45. https://doi.org/10.17762/ijritcc.v11i3s.6153

[43] Makarand L, M. . (2021). Earlier Detection of Gastric Cancer Using Augmented Deep Learning Techniques in Big Data with Medical Iot (Miot). Research Journal of Computer Systems and Engineering, 2(2), 22:26. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/28