An application of deep learning techniques for fraud detection in financial organizations: using a metaheuristic algorithm to uncover deception schemes and defalcation

A. J. Stagliano
Professor of Accounting
Erivan K. Haub School of Business
Saint Joseph's University
Philadelphia, PA, USA

Keywords
deep learning, financial fraud, honey badger algorithm

Abstract
This paper describes a new approach to financial fraud detection. It proposes, then tests with a simulated database of transactions, a metaheuristic algorithm based on two-layer deep learning. With a hybrid modification of the basic deep learning model that applies a honey badger modification, demonstrably higher quality outcomes were obtained than with competing models.

I. Introduction

As a consequence of the sensitive nature of financial data and intricate billing systems, financial organizations have become a common target for digital fraud. To mitigate financial losses resulting from fraudulent activity, it is crucial to continuously enhance fraud detection systems as electronic payments become the norm. This study applies a new hybrid form of deep learning to assist in both detecting fraud and, ultimately, fending off systems breaches.

Mitigation of cyber-attack data systems penetration is essential for continued/improved maintenance of high-level financial data integrity. This current research work is motivated by the urgent need to counter increasingly sophisticated fraudulent activities that exploit "cracks" and "holes" in processing platforms and applications so as to illegally access sensitive credit card information and perpetrate other transaction-related fraud activity.

II. Background

Machine learning algorithms and predictive models enable automated analysis of vast datasets and the transactions that make inquiry to the contained data [1]. With deep-learning systems in place, real-time detection of data breaches and system violation can be achieved. In recent years, deep learning has emerged as a powerful subset of machine learning that excels in handling complex and unstructured data. Deep learning models, such as neural networks, can automatically discover hierarchical representations from raw data [2]. This ability to uncover patterns can enable quite accurate and sophisticated fraud detection.

As applied here, the fraud detection model includes three distinct phases: thorough data pre-processing, advanced feature extraction, and a two-layer deep learning metaheuristic algorithm. The actual model proposed and tested is informed by both convoluted neural network (CNN) optimization and radial basic function network (RBFN) tessellation capturing. The innovation tested is inclusion of a honey badger algorithm (HBA) to identify patterns for feature selection [3]. By deploying this model, the aim is to proactively find and mitigate financial fraud. A successful model could substantially reinforce the security of extant electronic transactions processing schema.

III. Machine learning

Selection of a deep learning modality is motivated by several considerations—and this model choice reflects both advantages and potential limitations of its application. The decision to focus on deep learning models arises, in the first instance, from an increasing prominence of these techniques in addressing complex and high-dimensional data [4]. Credit card transactions, for example, certainly meet these criteria.

Deep learning models, particularly those infused with CNN and RBFN, have demonstrated a remarkably adept capability for pattern recognition [5]. This quality makes them a primary candidate for detecting intricately contrived fraudulent activities. Moreover, deep learning models have shown the potential to automatically learn tiered categorical features and patterns from raw data [6]. This capability is especially useful and advantageous when dealing with intricate and evolving fraud patterns that might be challenging for traditional models, such as logistic regression or random forest searching, to capture.

However, it is appropriate to acknowledge that the judgment to omit these traditional detection models as generators of performance benchmarks could indeed limit any comprehensive evaluation of the proposed hybrid model's outcome attainment. It is well known that both logistic regression analysis and random forest search models offer both simple interpretability and observable transparency, and these traits can be crucial in sensitive domains like financial fraud detection [7].

HBA, as used here in a self-improvement learning mode, offers the tremendous advantage of efficient and effective feature selection [8]. It employs a sophisticated approach to identify the most relevant and informative features from a dataset and, thus, reduces dimensionality and enhances a model's ability to distinguish between genuine and fraudulent transactions, for example [9]. This algorithm's selection process contributes to improved model accuracy, faster inherent training times, heightened overall performance, and higher quality outputs.

Two-layer deep learning is employed as a classifier to combine optimized CNN and RBFN to capture several modeling advantages. CNNs excel at learning hierarchical features from data (particularly images). RBFNs are proficient in capturing complex patterns and a variety of non-linear relationships that exist in datasets [10]. The fusing of these two distinct, yet complementary, architectures, creates a classifier that effectively leverages the strengths of both models. The optimized CNN, lensed through HBA, further enhances feature extraction, thereby enabling the classifier to accurately identify subtle fraudulent behavior that occurs within online financial transactions. This hybrid approach significantly improves fraud detection accuracy and adds substantial robustness to the overall system.

IV. **Test data development**

Here is how the data were developed for assessment of the proposed fraud-detection algorithm assessment in this research project. The first dataset used is a robust collection of credit card fraud detection elements. It is a simulated dataset of credit card transactions from January 1 through December 31, 2022, that includes both legitimate and fraudulent transactions. The computer-generated dataset has transactions that came from a pool of 800 businesses using the credit cards of 1,000 customers.

The second dataset is the online payment fraud detection data base. The file contains several characteristics or attributes that describe each transaction, including the financial amount, time of day, location, device-generating information, and user profile. In addition, labels indicating whether or not a transaction is fraudulent are included in this dataset. These legitimacy designations may be made in response to several variables, including whether the user or financial institution identified the transaction as fraudulent or whether it displayed particular patterns, features, or behaviors that are linked frequently to fraudulent activity.

V. **Method of study**

Testing of the hybrid detection-of-fraud model was carried out as follows. First, the original data were prepared for testing. The acquired raw data were pre-processed through a data-cleaning algorithm to fix/remove incorrect, duplicate, or incomplete data in the dataset. Next, basic features of the dataset were computed. These were: central tendency, degree of dispersion, principal component analysis outcomes, and certain mutual information-based features.

The third step in preparing the data for testing was a feature fusion transformation. In this phase, the extracted features were joined together via score-level fusion to create a comprehensive feature set as the fundamental input to the fraud detection model. At this point, the optimized HBA technique was applied to the data developed in the previous step so that those features that were fused as the most likely evidences of fraudulent occurrences/behaviors were identified.

Finally, the two-layer deep learning technique was implemented, as a hybrid classifier, to train the algorithm using the fused feature set and learn patterns and characteristics of fraudulent behavior [11]. This application method of deep learning as a classifier combines the optimized convolutional neural network and radial basic function network as operatives in the analysis. CNN, when applied here, is optimized using the honey badger algorithm since it is deemed to be an efficient parsing dichotomization tool.

The eventual result from the multi-phase analysis is the detected outcome which is used to flag potential frauds in the financial transaction database.

VI. Results

A large, varied database—covering a relatively significant time period given the types of transactions of interest—was used to test the proposed fraud detection model. Have we created, with this proposed model, an improved methodological application as compared to those traditional applications currently in use? To make that determination, we compare the outcomes of our proposed analytic scheme against five other techniques often used for determining fraudulent transactions in our simulated database of financial transactions: (1) convoluted neural network, (2) radial basic function network, (3) octree-based convoluted neural network, (4) deep neural network, and (5) recurrent neural network.

For every testing routine we use the same database and deliver, for comparison, outcomes on seven specific metrics that can be used to describe the quality of each fraud detection identifier. In this context, "quality" is taken to mean the ability of the model to correctly establish whether the transaction of interest is in fact fraudulent. The measures are: (1) modality sensitivity, (2) outcome identification specificity, (3) accuracy of identification, (4) model precision, (5) model robustness F-statistic, (6) negative predictive value indicator, and (7) false positive rate.

Comparative Measures of Model Identification Quality

	2LDL	CNN	RBFN	O-BCNN	DNN	RNN
Sensitivity	0.99	0.94	0.89	0.95	0.89	0.88
Specificity	0.97	0.95	0.89	0.97	0.89	0.88
Accuracy	0.98	0.95	0.91	0.96	0.91	0.89
Precision	0.98	0.96	0.91	0.97	0.90	0.89
F-score	9.96	8.44	9.33	9.52	9.93	9.00
NPV	0.98	0.95	0.91	0.96	0.91	0.90
FPR	0.98	0.94	0.92	0.94	0.91	0.90

2LDL = Two-layer deep learning
O-BCNN = Octree-based neural network
DNN = Deep neural network
RNN = Recurrent neural network
NPV = Negative predictive value
FPR = False positive rate

The table above indicates the clear domination of our proposed model in each category of identification quality. The proposed model, as tested with the simulated dataset, is a better fraud-detection instrument than those against which it was judged. Two other indicators that are additional signs of usefulness of the proposed mechanism for fraud detection in financial transactions are: (1) a very minimal 3 percent outcome for the full dataset find-and-replace measure, and (2) nearly 97 percent of outcomes occurring under the identification curve.

VII. Conclusion

This research project was undertaken to demonstrate that non-traditional computerized fraud detection models can be useful in stifling cybercrime in the financial services industry. In particular, suspected financial-transaction fraud is detectable in real time even in the most complex and convoluted

datasets. Employment of a two-layer deep learning metaheuristic algorithm can facilitate prediction and identification of both traits and patterns of fraudulent activity. The optimized honey badger routine applied here has wide-ranging application to problems, like fraud identification, in complex search-space that will benefit from enhanced pattern/feature identification convergence speed. Financial transaction malfeasance is both early-determinable and preventable. This research demonstrates a useful tool in that regard.

References

T. Hastie, R. Tibshirani, and J. H. Friedman, The Elements of Statistical Learning. New York: Springer, 2009.

A. I. Canhoto and F. Clear, "Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential," Business Horizons, vol. 63, pp. 183-193, 2020.

F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk, and W. Al-Atabany, "Honey badger algorithm: New metaheuristic algorithm for solving optimization problems," Mathematics and Computers in Simulation, vol. 192, pp. 84-110, February 2022.

Z-H. Zhou, Ensemble Learning Methods: Foundations and Algorithms, Boca Raton, FL: CRC Press, 2012.

Y. Bao, B. Ke, B. Li, J. Yu, and J. Zhang, "Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach," Journal of Accounting Research, vol. 58, no. 1, pp. 199-235, March 2020.

M. Fernandez-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?" Journal of Machine Learning Research, vol. 15, pp. 3133-3181, 2014.

J. L. Perols, R. M. Bowen, C. Zimmermann, and B. Samba, "Finding needles in a haystack: Using data analytics to improve fraud detection," The Accounting Review, vol. 92, pp. 221-245, 2017.

A. J. Stagliano and G. J. Tanzola, "Disrupting the accounting and financial reporting functions with implementation of artificial intelligence applications," unpublished manuscript presented at The Global Interdisciplinary Green Cities Conference 2023, University of Augsburg, June 2023.

MIT SMR Connections, How AI Changes the Rules: New Imperatives for the Intelligent Organization. Cambridge, MA: MIT Press, 2023.

E. Tuv, A. Borisov, G. Runger, and K. Torkkola, "Feature selection with ensembles, artificial variables, and redundancy elimination," The Journal of Machine Learning Research, vol. 10, pp. 1341-1366, 2009.

T. M. Khoshgoftaar, J, Van Hulse, and A. Napolitano, "Comparing boosting and bagging techniques with noisy and imbalanced data," IEEE Transactions on Systems, Man, and Cybernetics: Part A—Systems and Humans, vol. 4, pp. 552-568, 2011.