Implementing Deep Learning in E-Commerce Platforms for Fraud Detection and Management

Mary Jane C. Samonte
Doctor in Information Technology
Faculty Member of School of Information Technology
Mapua University, Makati City, Philippines
mjcsamonte@yahoo.com

Elcid A. Serrano
Doctor in Information Technology
Faculty Member of School of Information Technology
Mapua University, Makati City, Philippines
easerrano@mapua.edu.ph

Joseph Anthony T. Arpilleda, Robert Leyton-Pete J. Pastrana
Dhan L. Quijano, Romeo Robert R. Sulit
School of Information Technology, Mapua University
Makati City, Philippines, jatarpilleda@mymail.mapua.edu.ph,
rljpastrana@mymail.mapua.edu.ph, dlquijano@mymail.mapua.edu.ph,
rrrsulit@mymail.mapua.edu.ph

Abstract

E-Commerce has quickly become an essential facet of business transactions within the past decade as online platforms have paved the way towards convenient shopping that does not require the consumer to invest time and resources in traveling to the physical location of businesses. Along with the utilization of digital platforms to exchange goods, this also introduced the integration of digital payment methods such as e-wallets and credit cards. The use of fraud detection systems seeks to accurately filter and detect any malicious transactions that use unverified or illegitimate payment details. This paper aims to conduct a systematic literature review that will analyze collected academic papers from 2017 to 2021, which focus and relate to various deep learning techniques used in model creation. The findings from the systematic review will be used to develop and conceptualize a system that will utilize an efficient deep-learning method towards accurate fraud detection.

Keywords
E-Commerce, Fraud Detection System, Deep Learning

1. Introduction

Fraud detection systems are implemented to ensure the validity and legitimacy of transactions. Implementing Machine Learning algorithms has allowed for efficient systems that allow for accurate and real-time data monitoring (Abakarim et al, 2018). This is in line with the emergence of digital platforms for business, better known as E-Commerce, where goods and services are offered and advertised through online services (Chen and Lai, 2021). Fraudulent transactions have often been rooted in the use of credit cards, as most enterprises provide it as a payment method. The increase of online commerce has allowed for more significant instances of fraud, thus the need for greater systems that can detect and identify these cases (Asha and KR, 2021).

The researches that focus on the creation of fraud detection systems have often varied with the machine learning technique utilized, such as ones that use modified neural networks (Arya and Sastry, 2020) (Can et al, 2020) (Chen and Lai, 2021) (Gómez et al, 2018)(Kim et al, 2019) as well as auto-encoders (Gomes et al, 2018)(Liu et al 2020)

(Pumsirirat and Yan, 2018) in creating their respective models. The amount of research that delves into utilizing deep learning in fraud detection has increased within the five-year timeframe from 2017 to 2021.

This paper seeks to tackle and present the algorithms for data creation and consider the datasets for these models through their usage in the fraud detection systems discussed in the related literature.

2. Methodology

This study made use of the following online databases: Research Gate (researchgate.net), Science Direct (sciencedirect.com), IEEE Xplore (ieeexplore.ieee.org), Jurnal Ahmar (jurnal.ahmar.id), Hindawi (hindawi.com), IRO Journals (irojournals.com), and Semantic Scholar (semanticscholar.org)

2.1 Study Filtering

The selection process begins with a total of forty (40) studies gathered. In Table 1, it shows the objective, methods, and results of the studies. After filtering based on research title and abstract, there are thirty-two (32) studies left. The selected and screened studies are Research Gate (8), IEEE (5), Science Direct (4), Semantic Journals (1), IRO Journals (1), Hindawi (1), and Jurnal Ahmar (1), with a total of twenty-one (21) studies. The studies are selected and filtered based on the created criteria: the study uses deep learning techniques for fraud detection. Table 1 shows the screened studies and is arranged based on the number of pages of the study.

Table 1. Studies selected after undergoing selection.

Title of Study	Source/Publisher	Deep Learning Technique/s Used	Number of Pages
Deep learning for detecting financial statement fraud (Craja et al, 2020)	Science Direct	Neural Embeddings, Deep Learning, Hierarchical Attention Network	46
Insurance fraud detection with unsupervised, deep learning (Gomes et al, 2018)	Research Gate	Autoencoder, variational Autoencoder	30
Champion-challenger analysis for credit card fraud detection: hybrid ensemble and deep learning (Kim et al, 2019)	Science Direct	Convolutional Neural Network, Recurrent Neural Network	26
Doing good by fighting fraud: Ethical anti-fraud systems for mobile payments (Abi Din, et al, 2021)	Semantic Scholar	Generative Adversarial Network (GAN)	18
Predicting online shopping behavior from clickstream data using deep learning (Koehn et al, 2020)	Research Gate	Feature Engineering	16
A Closer Look Into the Characteristics of Fraudulent Card Transactions (Can et al, 2020)	IEEE	Multi-Layer Perceptron. Naive Bayes, Decision Tree, Random Forest,	15
Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata (Severino and Peng, 2021)	Science Direct	Deep Neural Network, Ensemble Model	14
DEAL – 'Deep Ensemble Algorithm' Framework for Credit Card Fraud Detection in	Research Gate	Extra-Tree Ensemble, Deep Neural Network	13

Real-Time Data with Google TensorFlow (Arya and Sastry, 2020)			
Utilizing Unlabeled Data to Detect Electricity Fraud in AMI: A Semisupervised Deep Learning Approach (Hu et al, 2019)	IEEE	The multitask feature extracting fraud detector a deep-learning-based model (MFEFD	13
Deep Convolution Neural Network Model for Credit-Card Fraud Detection and Alert (Chen and Lai, 2021)	IRO Journals	Deep Convolution Neural Network	12
Kollector: Detecting Fraudulent Activities on Mobile Devices Using Deep Learning (Sun et al, 2020)	IEEE	Multi-View Bagging Deep Learning	12
Session-Based Fraud Detection in Online E-Commerce Transactions Using Recurrent Neural Networks (Wang et al, 2017)	Research Gate	Recurrent Neural Networks	12
Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection (Oblé and Bontempi, 2019)	Research Gate	Deep Neural Network, BDNN, NDNN, FEDADNN, AugDNN, AdvNN	11
Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data (Liu et al 2020)	Hindawi	Stacked Denoising Auto-Encoder	11
End-to-end Neural Network Architecture for Fraud Scoring in Card Payments (Gómez et al, 2018)	Research Gate	Artificial Neural Networks	11
Spectral-Cluster Framework for Credit-Card Fraud Detection Using A Genetic Algorithm Trained Modular Deep Learning Neural Network (Ojugo and Nwankwo, 2021)	Jurnal Ahmar	Modular Neural Network	10
Privacy Preserving Distributed Deep Learning and Its Application in Credit Card Fraud Detection (Wang et al, 2018)	IEEE	Distributed Deep Learning	9
Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine (Pumsirirat and Yan, 2018)	Research Gate	Auto-Encoder, Keros Neural Network	8
Deep fraud. A fraud intention recognition framework in public transport context using a deep-learning approach (Herrera et al, 2018)	IEEE	Deep neural network	8
Credit card fraud detection using artificial neural network (Asha and KR, 2021)	ScienceDirect	Artificial Neural Network	7
An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning (Abakarim et al, 2018)	Research Gate	Auto-Encoder Deep Neural Network	7

3. Results and Discussion

This paper seeks to discuss and analyze the results that were obtained by the selected references. These include the following: (1) The focus and application of the fraud detection systems. (2) The type of deep learning technique that was utilized by each study. And (3) the results obtained from the models, such as accuracy and detection rate.

3.1 Fraud Detection

The study of fraud detection continues to grow as fraudulent transactions become mainstream, thus harming the industry. The use of fraud detection has been proven effective in preventing fraudulent transactions from happening. Fraud detections are essential in a business, especially when dealing with transactions. The twenty-one (21) studies show the focus of the study in creating fraud detection systems, as enumerated in Table 2.

Table 2. The focus of the selected reference studies.

Focus of Study	Number of Papers	Reference Studies
Card Transactions	12	(Arya and Sastry, 2020), (Can et al, 2020), (Chen and Lai, 2021), (Abi Din, et al, 2021), (Gómez et al, 2018), (Kim et al, 2019), (Ojugo and Nwankwo, 2021), (Pumsirirat and Yan, 2018), (Asha and KR, 2021), (Sun et al, 2020), (Wang et al, 2018), (Abakarim et al, 2018)
Online E-Commerce	3	(Koehn et al, 2020), (Liu et al, 2020), (Wang et al, 2017)
Financial Statements	1	(Craja et al, 2020)
Electricity	1	(Hu et al, 2019)
Property Insurance	1	(Severino and Peng, 2021)
Public transport	1	(Herrera et al, 2018)
Mobile Device	1	(Sun et al, 2020)
Insurance	1	(Gomes et al, 2018)

From the table above, some studies are similar when discussing the focus of the study, such as property insurance and insurance. The majority of the study focuses on the card transaction in developing an effective fraud detection system

3.2 Deep Learning Technique used

The deep learning techniques and models used in the selected studies are used in developing a fraud detection system. The use of deep learning techniques allowed the selected studies to create and develop an improved version of a fraud detection system. There are thirty-five (35) deep learning techniques that were used in the selected studies. These are the following techniques that are above 5% usage, the most used deep learning technique is Neural Networks with 51.43% usage, followed by Auto Encoder with 14.29% usage, and the last is the Ensemble Model with 5.71% usage, the rest of the deep learning techniques are only used once with 2.86% usage.

Table 3. Deep Learning techniques utilized by selected studies

Deep Learning Technique Used	Reference Studies
Neural Networks	(Arya and Sastry, 2020), (Can et al, 2020), (Chen and Lai, 2021), (Gómez et al, 2018), (Kim et al, 2019), (Oblé and Bontempi, 2019),

	(Herrera et al, 2018), (Ojugo and Nwankwo, 2021), (Pumsirirat and Yan, 2018), (Asha and KR, 2021), (Severino and Peng, 2021), (Wang et al, 2017), (Abakarim et al, 2018)
Auto Encoder	(Gomes et al, 2018), (Liu et al 2020), (Pumsirirat and Yan, 2018), (Abakarim et al, 2018)
Ensemble Model	(Arya and Sastry, 2020), (Severino and Peng, 2021)

Neural Network - This is the most utilized deep learning technique among all studies used in this paper. It is utilized by thirteen (13) studies. These networks are systems that use interconnected computing units or nodes that are based on the structure of how the human brain works (Arya and Sastry, 2020).

Auto Encoder - This is the second-most utilized deep learning technique in four (4) papers. Autoencoders are mainly used in unsupervised learning that uses the encoding and decoding of data to represent the data that was inputted (Pumsirirat and Yan, 2018).

Ensemble Models - The third-most utilized the deep learning technique used in two (2) research papers from the ones selected. Ensemble models are where various models are used to create the prediction; therefore, by combining multiple models, the user may increase the total accuracy of the predictions being made (Severino and Peng, 2021).

3.3 Results Observed

The results of the studies are reviewed adequately by observing the accuracy of the datasets provided from the studies used in testing the proposed algorithm or model from the studies. All of the studies showed that the developed algorithms or models used in testing the datasets had shown improvement over time. The developed algorithms or models from the selected studies can detect better accuracy than other models. The majority of the studies achieved 90% and higher accuracy when testing the model or algorithm from the datasets. Some results of the studies did not reach 90% or higher in accuracy but have shown improvement over more extended usage.

4. Proposed Fraud Detection

The proposed fraud detection system utilizes an artificial neural network as the deep learning algorithm for model creation. It has produced high accuracy rates in most studies that have been used. This type of system is said to bring higher accuracy rates than systems that use unsupervised algorithms such as auto-encoders (Asha and KR, 2021). Using Tensorflow, which uses neural networks and deep learning algorithms (Abakarim et al, 2018), a model shall be created from datasets that contain a vast array of data, such as in study (Pumsirirat and Yan, 2018) where the three datasets used have varying instances with instances that are 1000, 690, and 284,807 transactions respectively.

Once the model is trained and created, the testing of its accuracy shall be through binary analysis, which gives out a result of either 0 or 1 depending on whether or not a fraudulent transaction was detected (Asha and KR, 2021). To further enhance the system's accuracy, the use of Ensemble models may be utilized as they may bring significant improvements in classification, as highlighted by study (Arya and Sastry, 2020), where the neural network combined with an Ensemble algorithm brought a detection rate close to 100% accuracy. Figure 1 represents a graphical representation of the proposed system.

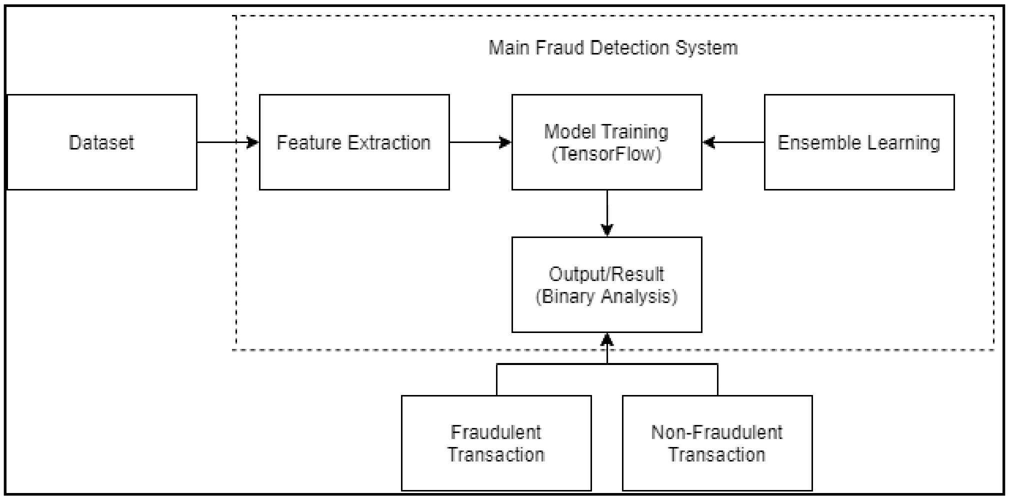
Main Fraud Detection System
Dataset
Feature Extraction
Model Training (TensorFlow)
Ensemble Learning
Output/Result (Binary Analysis)
Fraudulent Transaction
Non-Fraudulent Transaction

Figure 1. Proposed Fraud Detection System using ANN and Ensemble Model.

5. Implication and Conclusion

With the e-commerce industry rapidly developing and being more utilized, a detection system that determines whether a transaction is fraudulent or not is essential. E-commerce companies can implement and use the proposed model to identify transactions deemed fraud by the system based on the similarities of previous fraudulent transactions. The implications are presented using the data gathered and the data of related studies.

5.1 Practical Implication

Correct transactions are vital in satisfaction and providing better services to the consumers of an e-commerce platform. Because fraudulent transactions often use other people's identities, it can pose a severe threat to the privacy of people who got their identities stolen, leaving their identities vulnerable. The proposed system can stop the transaction from proceeding, preventing the breach of confidentiality from occurring. Implementing the fraud detection system could improve consumer satisfaction as their personal information is kept safe from being misused. An effective fraud detection system is beneficial for the employees of an e-commerce company, as they are the people that take responsibility for instances of fraudulent transactions. A fraud detection system can eliminate the fear of large and suspicious transactions, increasing employees' productivity and better engagements with e-commerce platforms.

5.2 Theoretical Implication

Previous research identified how fraud detection systems operate in preventing occurrences of fraudulent transactions. It was proven that fraud detections systems are essential in commerce, especially those that are conducted online. However, those researches focused on card transactions. E-commerce does not only use cards for transactions but other digital financial technology services. Ensuring that e-commerce is the field concentrated, the study and development of the proposed model provides a detailed explanation of behavioral similarities, additional information in online banking, and using deep learning in constructing detection systems, which is significant in the framework followed for this study.

The increase in fraudulent transactions in e-commerce platforms is prevalent, rising from 43% to 48% in 2020. Fraud detection systems are significant in maintaining satisfactory and efficient delivery of services in e-commerce platforms. By understanding patterns and the current situation of e-commerce in dealing with fraudulent transactions, this study proposes a model that identifies fraudulent transactions using deep learning. Through the data of previous studies and similarities, the effectiveness is verified. The verification proves that using systems specializing in fraud detection is possible and is beneficial for e-commerce. If implemented, deep learning in fraud detection systems for e-commerce will serve as helpful in increasing the efficiency, reliability, and improvement in the experience of the utilization of e-commerce.

6. Limitation and Future Research

This research is limited to the Fraud Detection of E-Commerce only, meaning that we do not be tackle other frauds that physical stores do since the definition of E-Commerce refers to online transactions. Since this study uses Neural Networks, Auto Encoder, and Ensemble Models, future researchers may use other deep learning techniques such as Boltzmann Machines to explore new ways of detecting fraud and test its accuracy in detecting it. Since this study focuses on Fraud Detection, one may be able to implement or conduct a study regarding countermeasures against it for future research. This study also recommends the integration and combination of various models and algorithms to determine effective combinations that will improve the efficiency and overall accuracy of the fraud detection system.

References

Abi Din, Z., Venugopalan, H., Lin, H., Wushensky, A., Liu, S. and King, S.T., Doing good by fighting fraud: Ethical anti-fraud systems for mobile payments. In *2021 IEEE Symposium on Security and Privacy (SP)* (pp. 1623-1640). IEEE, May 2021.

Abakarim, Y., Lahby, M. and Attioui, A., An efficient real time model for credit card fraud detection based on deep learning. In *Proceedings of the 12th international conference on intelligent systems: theories and applications* (pp. 1-7), October 2018.

Arya, M. and Sastry G, H., DEAL–'Deep Ensemble ALgorithm'framework for credit card fraud detection in real-time data stream with Google TensorFlow. *Smart Science*, *8*(2), pp.71-83, 2020.

Asha, R.B. and KR, S.K., 2021. Credit card fraud detection using artificial neural network. *Global Transitions Proceedings*, vol. *2*(1), pp.35-41, 2021.

Can, B., Yavuz, A.G., Karsligil, E.M. and Guvensan, M.A., A Closer Look Into the Characteristics of Fraudulent Card Transactions. *IEEE Access*, vol. *8*, pp.166095-166109, 2020.

Chen, J. and Lai, K., Deep Convolution Neural Network Model for Credit-Card Fraud Detection and Alert. June 2021, vol. 3(2), pp.101-112, 2021.

Craja, P., Kim, A. and Lessmann, S., Deep learning for detecting financial statement fraud. *Decision Support Systems*, vol. 139, p.113421, 2020.

Gomes, C., Jin, Z. and Yang, H., Insurance fraud detection with unsupervised deep learning. *Journal of Risk and Insurance*, vol. 88(3), pp.591-624, 2021.

Gómez, J.A., Arévalo, J., Paredes, R. and Nin, J., End-to-end neural network architecture for fraud scoring in card payments. *Pattern Recognition Letters*, vol 105, pp.175-181, 2018.

Herrera, J.L.L., Figueroa, H.V.R. and Ramírez, E.J.R., February. Deep fraud. A fraud intention recognition framework in public transport context using a deep-learning approach. In *2018 International Conference on Electronics, Communications and Computers (CONIELECOMP)* (pp. 118-125). IEEE, 2018.

Hu, T., Guo, Q., Shen, X., Sun, H., Wu, R. and Xi, H., Utilizing unlabeled data to detect electricity fraud in AMI: A semisupervised deep learning approach. *IEEE transactions on neural networks and learning systems*, vol. 30(11), pp.3287-3299, 2019.

Kim, E., Lee, J., Shin, H., Yang, H., Cho, S., Nam, S.K., Song, Y., Yoon, J.A. and Kim, J.I., Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning. *Expert Systems with Applications*, vol. 128, pp.214-224, 2019.

Koehn, D., Lessmann, S. and Schaal, M., Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, vol.150, p.113342, 2020.

Liu, J., Gu, X. and Shang, C., 2020. Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data. *Complexity*, 2020.

Oblé, F. and Bontempi, G., Deep-learning domain adaptation techniques for credit cards fraud detection. In *Recent Advances in Big Data and Deep Learning: Proceedings of the INNS Big Data and Deep Learning Conference* (Vol. 1, pp. 78-88), April 2019.

Ojugo, A.A. and Nwankwo, O., Spectral-Cluster Solution For Credit-Card Fraud Detection Using A Genetic Algorithm Trained Modular Deep Learning Neural Network. *JINAV: Journal of Information and Visualization*, vol. 2(1), pp.15-24, 2021.

Pumsirirat, A. and Yan, L., Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine. *International Journal of advanced computer science and applications*, vol. 9(1), pp.18-25, 2018.

Severino, M.K. and Peng, Y., Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata. *Machine Learning with Applications*, p.100074, 2021.

Sun, L., Cao, B., Wang, J., Srisa-an, W., Philip, S.Y., Leow, A.D. and Checkoway, S., Kollector: Detecting fraudulent activities on mobile devices using deep learning. *IEEE Transactions on Mobile Computing*, vol. 20(4), pp.1465-1476, 2020.

Wang, S., Liu, C., Gao, X., Qu, H. and Xu, W., September. Session-based fraud detection in online e-commerce transactions using recurrent neural networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 241-252). Springer, Cham, 2017.

Wang, Y., Adams, S., Beling, P., Greenspan, S., Rajagopalan, S., Velez-Rojas, M., Mankovski, S., Boker, S. and Brown, D., August. Privacy preserving distributed deep learning and its application in credit card fraud detection. In *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (pp. 1070-1078). IEEE, 2018.

Biographies

Mary Jane C. Samonte has a double bachelor's degree in computer education and information technology. She also has two post graduate degree; Information Technology and Computer Science. She finished her Doctor in IT with a study focusing in Deep Learning. She has a wide range of research interests that are centered around educational technologies, gamification, mobile and ubiquitous learning, digital game-based learning, artificial intelligence in education, e-health, assistive technology, natural language processing, green computing and data analytics-based studies.

Elcid A. Serrano holds a Doctor of Information Technology degree with research interests in software engineering, human computer interaction, IT-enabled services, and grounded theory in information systems.

Joseph Anthony T. Arpilleda is a second-year Bachelor of Science in Information Technology student of Mapua University.

Robert Leyton-Pete J. Pastrana is a second-year Bachelor of Science in Information Technology student of Mapua University.

Dhan L. Quijano is a second-year Bachelor of Science in Computer Science student of Mapua University.

Romeo Robert R. Sulit is a second-year Bachelor of Science in Information Technology student of Mapua University.