JOURNAL OF APPLIED INTELLIGENT SYSTEMS & INFORMATION SCIENCES

VOL. 3, ISSUE 2, PP. 65-73, DECEMBER 2022.
Available at: https://journal.research.fanap.com/
DOI: https://doi.org/10.22034/JAISIS.2022.377265.1057

ENSEMBLE LEARNING FOR FRAUD DETECTION IN E-COMMERCE TRANSACTIONS: A COMPARATIVE STUDY

Mohammad Amini [1,*], Mohammad Rabiei [2]

[1] *School of Information Technology, Mehralborz University, Tehran, Iran,*

[2] *Department of Information Technology, Iranian Research Institute for Information Science and Technology (IRANDOC), Tehran, Iran.*

ABSTRACT

Fraud detection is a critical measure in today's digitalized world and e-commerce platforms. Considering the significant advantages that ensemble methods bring to the world of machine learning, it seems necessary to examine their potential usage in the field of fraud detection. This is especially important for e-commerce transactions where we need to assess whether Ensemble methods can be good candidates as effective classifiers. In this paper, we evaluate the potential of ensemble learning methods for fraud detection in the e-commerce domain. We implement several well-known ensemble methods on e-commerce customer data and compare their result using different performance criteria. Our results show that *XGBoost* and *Random Forest* outperform other ensemble methods for fraud detection. The results of this study can be helpful for those scholars who are willing to optimize their fraud detection systems with ensemble methods. Also, the present study shows which classification algorithms can be best used in an ensemble framework to be applied in fraud detection for online payments.

***KEYWORDS*:** Fraud Detection, E-commerce, Ensemble Learning

a) INTRODUCTION

E-commerce has seen a lot of growth and development in recent years. Digital transformation and advanced e-commerce platforms along with the growth of communication capabilities have propelled consumers into more online shopping (Song et al., 2022). Businesses and customers can receive many benefits by using Ecommerce platforms; The buying process is faster, many costs are reduced and there is high flexibility for customers as the can compare the price and quality of the products, while they are offered several payment options (Rodrigues et al., 2022). Buying and selling goods online has grown even more due to the conditions of the COVID-19 pandemic, and it seems that this process is accelerating (Tran, 2021). According to recent surveys on eMarketer, worldwide e-commerce sales increased by 27.6% in 20202 and 14.3% in 2021 respectively (eMarketer Editors, 2021). Naturally, this rapid growth of online business platforms and the increase in the volume of transactions attract fraudsters and profit seekers and increase fraud cases. Fraudsters can gain huge sums of money from possible loopholes and vulnerabilities in electronic payment processes.

[*] *Corresponding Author, Email: amini@mehralborz.ac.ir*

RECEIVED: 18 DECEMBER, 2022; **ACCEPTED**: 30 DECEMBER, 2022; **PUBLISHED ONLINE**: 30 DECEMBER, 2022

Therefore, it is absolutely necessary to implement strong and smart fraud detection solutions to prevent these financial and economic losses.

Basic approaches used for fraud detection tend to analyze customer data and identify patterns associated with the fraud. The most important types of data investigated are online navigation tracks, historical activities, and the payment behavior of customers. Data mining and machine learning are among the most successful methods for fraud detection in domains such as credit card and financial fraud (Abdallah et al., 2016; Diadiushkin et al., 2019; Varmedja et al., 2019). Generally, the problem is formulated as a two-class classification in which any input transaction is labeled as normal or fraudulent. Machine learning methods can help solve this problem by learning from the train data and then applying the learned patterns to the production data. Popular classification algorithms such as Logistic Regression and K-nearest neighbor (Itoo & Singh, 2021), Artificial Neural Networks (Asha & KR, 2021), Support Vector Machines (Gyamfi & Abdulai, 2018), and Random Forests (Xuan et al., 2018) have been proposed for fraud detections so far. However, as the number and complexity of fraudulent attempts have increased in e-commerce transactions and the data for detection may have some problems including skewed and noisy data, it is difficult for these traditional methods to capture multiple characteristics as well as the underlying structure of data. Thus, we need more sophisticated and robust methods to deal with fraud in the modern age. Recent studies in machine learning applications for fraud detection have been more directed toward the use of hybrid and flexible systems (Lin et al., 2021; Nami & Shajari, 2018).

Ensemble methods are among the powerful solutions that can improve the accuracy of classification (Polikar, 2012). The main idea behind Ensemble Learning is to combine various classifiers with different learning mechanisms or training samples to improve the final prediction results (Dietterich, 2002; Dong et al., 2020). In other words, Ensemble Learning aims to integrate diversely supervised or unsupervised classification algorithms into a unified framework by using a combination method or voting system to improve the overall performance of the system. Fig. 1 shows a general architecture for an Ensemble Learning model that is based on supervised classification algorithms. To create an ensemble model, first, some different (usually weak) classifiers are trained on the training data to learn the patterns in data using their algorithm. Then the predictions obtained from each classifier are aggregated using a combination or voting method to produce a final prediction. Such a framework has many benefits for machine learning methods including extensibility for different methods, improving detection performance, and flexibility of usage (Kuncheva, 2014; Polikar, 2012). Therefore, ensemble methods have been used in a variety of applications including network intrusion detection (Amini et al., 2016), bio-informatics (Verma & Mehta, 2017), time-series forecasting (Galicia et al., 2019), and risk analysis (Hamori et al., 2018).

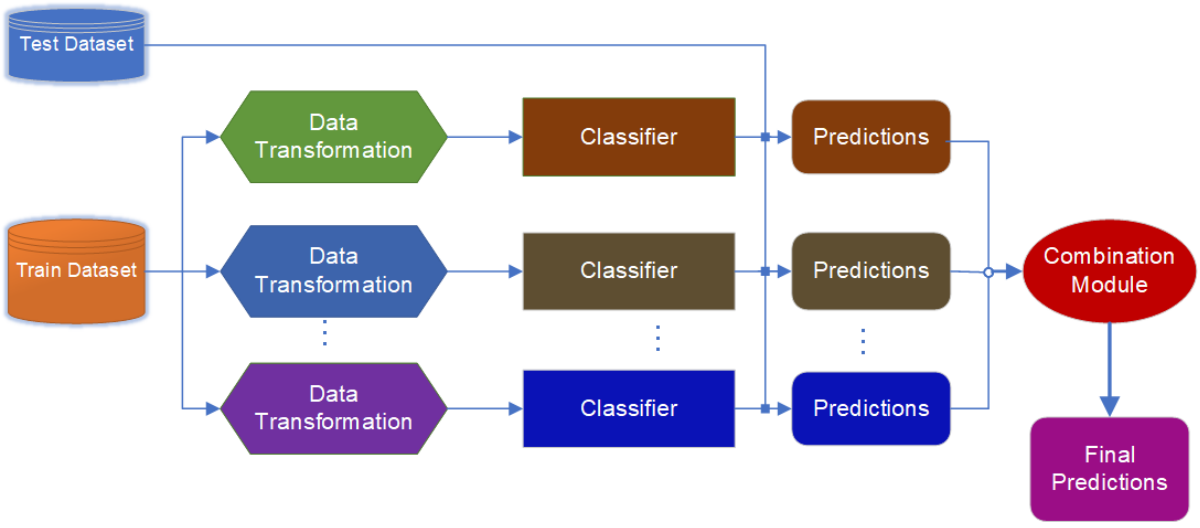

Fig. 1. A general framework for supervised Ensemble Learning

Considering the total benefits which ensemble methods provide us, it seems necessary to examine their application in the field of fraud detection, especially for e-commerce transactions, and assess whether they can be considered effective solutions. But until now, there has been no complete research on the possible improvement and effectiveness of these methods in fraud detection. In this paper, we evaluate the potential of Ensemble Learning methods to be used for fraud detection e-commerce domain. We implemented several well-known ensemble methods on the e-commerce customer data and compared their result using different performance criteria. The results of this study can be helpful for those scholars who are willing to optimize their fraud detection systems with ensemble methods. Also, the present study shows which classification algorithms can be best used in an ensemble framework for fraud detection in online payments.

The rest of this paper is organized as follows: in section 2, we provide a background on the methods proposed so far for fraud detection and an introduction to some ensemble methods' structures. Section 3 includes the presentation of our methodology for implementing different ensemble methods on the fraud data, the experimental procedure, and the evaluation phase. In Section 4, we present and analyze the evaluation result. Finally, in section 5 we conclude the paper and present some future directions.

b) RELATED WORK

In this section, we provide a background on Ensemble Learning systems and some of their applications in fraud detection.

a. Ensemble methods

During the last three decades, the popularity of ensemble methods has increased and different ensemble methods have been proposed by incorporating different aggregation, combination, or data sampling elements (Dietterich, 2002; Dong et al., 2020). The structure of the ensemble model can be defined by modification on four levels: the type of classifiers to be included, the sampling methods to be used, the feature set that is defined, and the combination module. Based on these levels, several ensemble methods have been proposed in the literature.

Bagging: This is the simplest form of an ensemble that is generated by a set of independent classifiers each of which is trained on a different sample taken from an original data set using sampling by replacement method (i.e. bootstrap sampling) (Breiman, 1996). This method ensures diversity in the ensemble by the variations within the bootstrapped samples on which each classifier is trained. A relatively weak classifier is used as a base classifier whose decision boundaries measurably vary concerning relatively small perturbations in the training data. To generate the final prediction for the new instances the Majority voting rule is performed for the individual prediction of the classifiers (Kuncheva, 2014).

AdaBoost: This method is one of the best-known ensemble methods used for many applications (Schapire, 2013). It uses the general boosting idea to develop the classifier ensemble incrementally, adding one classifier at a time. In AdaBoost, the focus is on instances that were previously misclassified when training a new classifier. The overall prediction performance is increased by implementing a weighing mechanism for instances and base classifiers. In each iteration, the weights of misclassified instances are increased, while the weights of correctly classified instances are decreased. In addition, weights are also assigned to the individual base learners based on their overall predictive performance. One important distinction in AdaBoost is that the training of base models is performed in a sequential manner instead of the parallel manner used in Bagging.

Random forest: This method creates an ensemble using a large number of decision trees that are unpruned and trained in a bagging framework (Breiman, 2001). In contrast to the split made in the decision tree classifier such as C4.5 (Safavian & Landgrebe, 1991), the decisions made in the trees in Random Forests incorporate randomness in the feature set used for splitting. In fact, instead of selecting the best attribute at each node (using, e.g., an information gain measure), the attribute is selected randomly in such a way that its probability of being selected is proportional to its measured value. Through applying several other ways to incorporate

randomness into a predictor besides the random forest procedure, some new forms and versions of ensembles were created.

Random Subspace: While Random forests have been built based on decision trees as base classifiers, in random subspace this assumption is removed. Random Subspace incorporates randomness into the ensemble by building a set of feature subspaces via randomly sampling features and then trains basic classifiers – of any kind- in these subspaces to generate multiple results (Ho, 1998).

Gradient Boosting Machines (GBM): This method is formed by computing a sequence of regression trees where each tree in the sequence predicts the pseudo-residuals of the preceding trees given an arbitrary differentiable loss function (Natekin & Knoll, 2013). The aggregation of the Predictions is performed in an additive manner where each added model is trained to minimize the loss function. Similar to what is conducted in AdaBoost, individual classifiers in gradient boosting are trained successively. It is important to note that a GBM model usually has many shallow trees, as opposed to a random forest which has fewer (but deeper) trees. A scalable version of this algorithm called XGBoost (Chen & Guestrin, 2016) has gained a lot of popularity due to its novel algorithmic optimizations and refinements. In addition, XGBoost provides a regularization component to the loss function in GBM, aimed at creating ensembles that are simpler and more generative.

b. Ensemble Learning for fraud detection

Due to the many capabilities that ensemble methods provide in improving detection and classification performance, researchers have tried to use these methods in the field of fraud detection. Most of the proposed methods have used Ensemble Learning as part of the detection algorithm along with other components in a hybrid structure. Among the main ensemble methods, many researchers have been able to use Random Forest to detect fraud (Carneiro et al., 2017; Dornadula & Geetha, 2019; Rai & Dwivedi, 2020). This method is mentioned as one of the most successful methods in the literature specifically for credit card fraud detection. Sohony et al., (2018) proposed an ensemble method for credit card fraud detection based on a combination of random forest and neural network that used the advantages of both these methods for more accurate detection of normal instances, and fraud instances.

An ensemble-based method was proposed by Haider et al., (2018) for impression fraud detection in mobile advertising. The authors used bagging and boosting ensemble methods to classify each ad display, also called an impression, as fraudulent or non-fraudulent. They could achieve a high rate of accuracy and precision, as well as good recall. Xu et al., (2011) proposed a random rough subspace-based neural network ensemble method for insurance fraud detection. A rough set reduction was first employed to generate a set of reductions that could keep the consistency of data information. Then, the reductions were randomly selected to construct a subset of reductions. Finally, each of the selected reductions was used to train a neural network classifier in an ensemble framework. Bagga et al., (2020) applied the bagging method along with pipelining to improve the performance of credit card fraud detection.

Although the ensemble methods show a promising future for the world of fraud detection, their vast potential for improving detection performance is yet to be investigated more systematically. In the next section, we present our methodology to evaluate the performance of some important ensemble methods for fraud detection.

c) RESEARCH METHODOLOGY

Our goal in this study is to assess the potential of Ensemble Learning for fraud detection in e-commerce transactions. Therefore, we have selected some of the best-known ensemble methods in the literature to be implemented as classification systems for fraud detection. We also use a data set that contains information about online payment fraud (Kharwal, n.d.). The set of ensemble methods for this evaluation are *Bagging*, *AdaBoost*, *Random Forest*, *Random Subspace* as well as *XGBoost*. As we can use any type of classifier in *Random Subspace*, we chose two important classification algorithms in the random subspace ensemble framework. They are *Neural Networks* (NN) and *Support Vector Machines* (SVM). Therefore, there are six ensemble methods used in our evaluation phase.

Table 1 shows the features of online payment fraud used in this study. This data set has been obtained from the Kaggle community and can be used to evaluate different fraud detection solutions. The data set contains 6362620 records of historical information about customer transactions. The target variable that is to be predicted is 'isFraud' in which class 0 represents a normal transaction and class 1 represents a fraudulent transaction. Naturally, such a data set is imbalanced. We chose 70% of the data for training the models and used 30% for testing. Also, for the training phase, we used 5-fold cross-validation. We also optimized all methods using cross-validation and grid search methods to get the best set of hyperparameters needed to get the results. So, our results have been obtained from an optimized set of methods and we compare the best results received from the ensemble methods. Typically, a fraud detection system assigns a label to each transaction instance that shows whether it is fraud or normal. The underlying implementation of most algorithms actually may assign a probability to each case that shows how confident the system is about the case is a fraud. This probability would be rounded to get a number in {0,1}.

We use a set of criteria for evaluating the performance and comparing the ensemble methods together. We consider four important detection metrics as True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP). True Positives are the instances that are positive and were also classified as positive. Similarly, True Negatives are the actual negatives and were classified as negative. False positives are cases that are negative but are classified as positives. Similarly, False Negatives are cases that are positive but are classified as negative. According to these metrics, the performance criteria are as follows:

Table 1. The online payment fraud data set features

Feature	Description
step	represents a unit of time where 1 step equals 1 hour
type	type of online transaction
amount	the amount of the transaction
nameOrig	customer starting the transaction
oldbalanceOrg	balance before the transaction
newbalanceOrig	balance after the transaction
nameDest	recipient of the transaction
oldbalanceDest	initial balance of the recipient before the transaction
newbalanceDest	the new balance of the recipient after the transaction
isFraud	Class label

Accuracy: The total number of instances that were classified correctly divided by the total number of instances. It can be shown below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision: The number of positive instances that were classified correctly divided by the total number of detected positive instances. It can be shown below:

$$Precession: \frac{TP}{TP + FP} \quad (2)$$

Recall: The number of positive instances that were classified correctly divided by the total number of actual positive instances. It can be shown bellow

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-Score: A combined metric that takes bot precision and criteria into consideration:

$$F\text{-}Score = \frac{2 \times Precesion \times Recall}{Precesion + Recall} \tag{4}$$

d) **RESULTS AND DISCUSSION**

After the experiments were completed, we extracted the results and organized them according to the performance criteria introduced in the previous section. Fig. 2 depicts the accuracy of six ensemble methods used in this study.

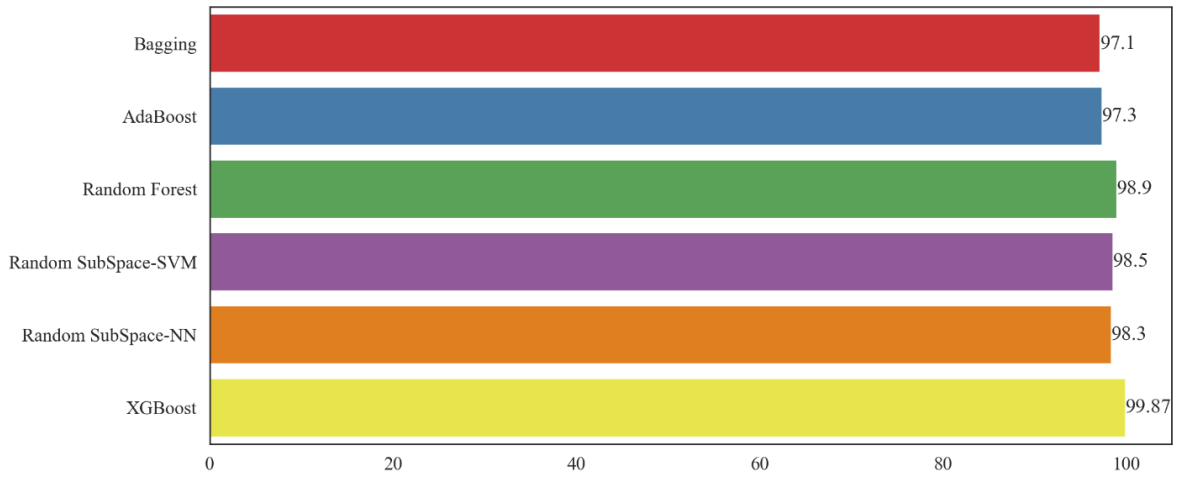

Fig. 2. Accuracy of six ensemble methods on online payment fraud data (%)

It can be seen that generally; all ensemble methods show good accuracy for detection. An accuracy of 97% can be generally considered a good performance in the fraud detection domain. Thus, we can conclude that ensemble methods are good promising methods that can be used as candidate solutions for fraud detection. However, *XGBoost* outperforms other methods with 99.8%. This is a great performance considering the accuracy metric.

Accuracy is a general metric that shows the overall classification performance of the system as far as the prediction for each instance is concerned. In the fraud domain where we have a class imbalance problem while having fewer false negatives is desired, accuracy cannot express the real detection performance of the methods. To have a better understanding of the ensemble methods' performance for the classification task, we need to make a comparison based on more specific criteria i.e., precession, recall, and F1-score. Table 2 shows a detailed evaluation of the performance metrics for two classes of the fraud data set. We calculated the classification metrics for both classes separately because it is very important to know how the classifier performs in normal and fraudulent cases. As can be seen in the table, *XGBoost* shows better performance on precession, recall, and F1-score.

For precession, it is obvious that all ensemble methods have gained almost 100% for class 0 (the normal transactions). That is because there is a high number of normal observations and the FP for class 0 is very low

compared to the TP. For class 1, however, the precession is very low for most of the methods showing that FP for class 1 is very high, i.e., a high number of normal cases have been classified as fraud. *XGBoost* shows an acceptable performance of 0.62 percent, which means it can prevail over the imbalance problem very well to get a good precession and a lower FP.

Class imbalance in fraud data is a major issue and is reflected in the classification criteria results. The number of fraud cases is much smaller than the number of normal records. In such cases, the model will be biased toward normal observations and may classify more fraud cases as normal leading to a high number of false negatives. Therefore, we need to consider recall as a more important metric in fraud classification. According to the recall in class 0, we can observe that all methods have gained high scores, with *XGBoost* having 1.00 for its recall. The reason for the high recall score in class 0 is the high number of normal records and the bias toward the normal cases in the model. *XGBoost* has been able to find all normal cases with no normal records predicted as fraud which show FN = 0 for class 0. For recall in class 1, the three ensemble methods *Bagging*, *AdaBoost*, and *XGBoost* show the best results, when Bagging surprisingly outperforms the two other methods with 92%. This simple ensemble method has predicted more fraud cases than other methods did. Thus, the traditional bagging and boosting methods can be good candidate classifiers if we want to emphasize more on recall more in fraud detection. Random forest and random subspace methods seem to provide lower recall for class 1.

Table 2. The performance comparison of six ensemble methods on online payment fraud data

Method	Precision		Recall		F1-Score	
	Class 0	***Class 1***	***Class 0***	***Class 1***	***Class 0***	***Class 1***
Bagging	1.00	0.12	0.96	0.92	0.98	0.21
Adaboost	1.00	0.14	0.99	0.91	0.99	0.24
Random Forest	1.00	0.46	0.98	0.88	0.99	0.6
Random SubSpace (SVM)	1.00	0.38	0.97	0.89	0.98	0.53
Random SubSpace (NN)	1.00	0.36	0.97	0.89	0.98	0.51
XGBoost	1.00	0.62	1.00	0.91	1.00	0.74

Considering F1-Score, while all ensemble methods provide high values for class 0, the key distinction emerges in class 1 (fraud class) where *XGBoost* provided the highest value and outperforms all ensemble methods by a significant distance. The most important thing to consider is that the highest amount of contribution of the large distance is provided by the high Precession of *XGBoost* where it can reduce FP for class 1 and predict the fraud cases more correctly than other methods. Considering that *XGBoost* could provide the best precession in class 1 and one of the best recalls in class 1, it can be chosen as the most suitable ensemble method for the fraud detection domain. *Random Forest*, which also provides a good score in precision, takes second place in F1-Score. The good performance of *Random Forest* is consistent with other studies which used Random Forest as an effective fraud detection method.

One important result that we can be obtained from this study is that *XGBoost* can be chosen as a promising classification algorithm to be used in ensemble systems for fraud detection. However, this selection is solely based on experimental evaluations. A suitable theoretical framework is needed to examine and recommend a single algorithm for use in the ensemble system. One needs to pay attention to the fact that our focus in this article was on the practical examination of Ensemble systems in the domain of e-commerce transactions, and we meant choosing a suitable algorithm for the Ensemble system only in this field. For this purpose, we compared different algorithms using different classification criteria.

e) CONCLUSION

Ensemble methods provide many benefits for classification performance. In this paper, we systematically evaluated the possibility of using popular ensemble methods applications in the field of fraud detection, especially for e-commerce transactions. We assessed whether they can be good candidates for effective solutions. We implemented several well-known ensemble methods on the e-commerce customer data and compared their results using different performance criteria. The results of this study showed that almost all ensemble methods perform well on fraud data. Among these ensemble methods, *Random Forest* and *XGBoost* can show superior performance and provide a good f1-score for the fraud class. The results of this study are consistent with those studies that introduce *Random Forest* as a promising method for fraud detection. Also, our investigation shows *Random Forest* and *XGBoost* can be best used in an ensemble framework for fraud detection in online payments.

Although our focus in this article was on the practical examination of Ensemble systems in the domain of e-commerce transactions, for future works, we will need to find a suitable theoretical framework to better evaluate the performance of multiple ensemble systems and find a more robust method to recommend a single algorithm for use in the ensemble system. We can also evaluate the ensemble methods on other domains such as banking fraud and insurance fraud to get more insights into the ensemble methods' performance.

REFERENCES

Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, *68*, 90–113.

Amini, M., Rezaeenour, J., & Hadavandi, E. (2016). A neural network ensemble classifier for effective intrusion detection using fuzzy clustering and radial basis function networks. *International Journal on Artificial Intelligence Tools*, *25*(02), 1550033.

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

Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). Credit card fraud detection using pipeling and ensemble learning. *Procedia Computer Science*, *173*, 104–112.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Carneiro, N., Figueira, G., & Costa, M. (2017). A data mining based system for credit-card fraud detection in e-tail. *Decision Support Systems*, *95*, 91–101.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794.

Diadiushkin, A., Sandkuhl, K., & Maiatin, A. (2019). Fraud detection in payments transactions: Overview of existing approaches and usage for instant payments. *Complex Systems Informatics and Modeling Quarterly*, *20*, 72–88.

Dietterich, T. G. (2002). Ensemble learning. *The Handbook of Brain Theory and Neural Networks*, *2*(1), 110–125.

Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, *14*(2), 241–258.

Dornadula, V. N., & Geetha, S. (2019). Credit card fraud detection using machine learning algorithms. *Procedia Computer Science*, *165*, 631–641.

eMarketer Editors. (2021). *Worldwide e-commerce will approach $5 trillion this year*. https://www.insiderintelligence.com/content/worldwide-ecommerce-will-approach-5-trillion-this-year

Galicia, A., Talavera-Llames, R., Troncoso, A., Koprinska, I., & Martínez-Álvarez, F. (2019). Multi-step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, *163*, 830–841.

Gyamfi, N. K., & Abdulai, J.-D. (2018). Bank fraud detection using support vector machine. *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 37–41.

Haider, C. M. R., Iqbal, A., Rahman, A. H., & Rahman, M. S. (2018). An ensemble learning based approach for impression fraud detection in mobile advertising. *Journal of Network and Computer Applications*, *112*, 126–141.

Hamori, S., Kawai, M., Kume, T., Murakami, Y., & Watanabe, C. (2018). Ensemble learning or deep learning? Application to default risk analysis. *Journal of Risk and Financial Management*, *11*(1), 12.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832–844.

Itoo, F., & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes, and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, *13*(4), 1503–1511.

Kharwal, A. (n.d.). *Online Payment Fraud Detection* [Kaggle]. https://www.kaggle.com/datasets/jainilcoder/online-payment-fraud-detection

Kuncheva, L. I. (2014). *Combining pattern classifiers: Methods and algorithms*. John Wiley & Sons.

Lin, W., Sun, L., Zhong, Q., Liu, C., Feng, J., Ao, X., & Yang, H. (2021). Online Credit Payment Fraud Detection via Structure-Aware Hierarchical Recurrent Neural Network. *IJCAI*, 3670–3676.

Nami, S., & Shajari, M. (2018). Cost-sensitive payment card fraud detection based on dynamic random forest and k-nearest neighbors. *Expert Systems with Applications*, *110*, 381–392.

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*, 21.

Polikar, R. (2012). Ensemble learning. In *Ensemble machine learning* (pp. 1–34). Springer.

Rai, A. K., & Dwivedi, R. K. (2020). Fraud detection in credit card data using machine learning techniques. *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences*, 369–382.

Rodrigues, V. F., Policarpo, L. M., da Silveira, D. E., da Rosa Righi, R., da Costa, C. A., Barbosa, J. L. V., Antunes, R. S., Scorsatto, R., & Arcot, T. (2022). Fraud detection and prevention in e-commerce: A systematic literature review. *Electronic Commerce Research and Applications*, 101207.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 660–674.

Schapire, R. E. (2013). Explaining AdaBoost. In *Empirical inference* (pp. 37–52). Springer.

Sohony, I., Pratap, R., & Nambiar, U. (2018). Ensemble learning for credit card fraud detection. *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, 289–294.

Song, Y., Escobar, O., Arzubiaga, U., & De Massis, A. (2022). The digital transformation of a traditional market into an entrepreneurial ecosystem. *Review of Managerial Science*, *16*(1), 65–88.

Tran, L. T. T. (2021). Managing the effectiveness of e-commerce platforms in a pandemic. *Journal of Retailing and Consumer Services*, *58*, 102287.

Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., & Anderla, A. (2019). Credit card fraud detection-machine learning methods. *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 1–5.

Verma, A., & Mehta, S. (2017). A comparative study of ensemble learning methods for classification in bioinformatics. *2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, 155–158.

Xu, W., Wang, S., Zhang, D., & Yang, B. (2011). Random rough subspace based neural network ensemble for insurance fraud detection. *2011 Fourth International Joint Conference on Computational Sciences and Optimization*, 1276–1280.

Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018). Random forest for credit card fraud detection. *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, 1–6.