Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 214 (2022) 117–124

Procedia Computer Science

www.elsevier.com/locate/procedia

9th International Conference on Information Technology and Quantitative Management

Exploratory analysis and implementation of machine learning techniques for predictive assessment of fraud in banking systems

Miguel Ângelo Lellis Moreira[a,b*], Claudio de Souza Rocha Junior[a],
Diogo Ferreira de Lima Silva[a], Marcos Alexandre Pinto de Castro Junior [c],
Igor Pinheiro de Araújo Costa[a,b], Carlos Francisco Simões Gomes[a],
Marcos dos Santos[b,c]

[a] *Fluminense Federal university, Niterói, RJ 24210-240, Brazil*
[b] *Naval Systems Analysis Center, Rio de Janeiro, RJ 20091-000, Brazil*
[c] *Brazilian Navy, Rio de Janeiro, RJ 20091-000, Brazil*
[d]*Military Institute of Engineering, Urca, RJ 22290-270, Brazil*

Abstract

Regarding combating financial fraud in banking systems, this study analyzes machine learning algorithms used in predictive assessments for fraud detection. For a given methodological approach, a database with more than six million records referring to the financial transactions of a given banking organization is used. First, an exploratory data analysis clarifies the main variables influencing the evaluation process, with binary and financial percentages related to fraud loss. Concerning the unbalance between the expected records and those classified as fraud, Random Under Sampling, SMOTE, and ADASYN techniques are used to balance and train these bases by Logistic Regression, Naive Bayes, KNN, and Perceptron techniques. With the implementation of machine learning algorithms, we present the main feasibilities of each model in each scenario. At the end of the study, we expose the final considerations and proposals for future work.

Keywords: Machine Learning, Big Data, and Bank Fraud.

1. Introduction

Technological advancement has increasingly influenced modern life, making the population migration of physical services to online platforms inevitable [1]. Regarding the use of financial services of banking organizations, it has become common to move money entirely through web applications [2]. However, along with the migration to online platforms, there is an exponential increase in financial transactions, bringing a greater need for cybersecurity and techniques for predicting fraud on organizational systems [3–5].

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: miguellellis@hotmail.com.

10.1016/j.procs.2022.11.156

As exposed in [6], financial fraud do not impact only on the banking organization but also everyday life as a whole, leading to the reduction of the level of trust between organizations, destabilizing financial operations, and indirectly increasing the cost of living [7, 8]. Even though organizations present high levels of and a wide variety of fraud prevention methods, frauds are often inevitable [9] once fraudsters quickly adapt to those advance in prevention technologies [10, 11].

In this scenario, machine learning (ML) techniques enable evaluation and prediction models as a way to aid in the classification of routine or fraudulent financial transactions [12]. As stated in [13], ML is understood as a data analysis method that provides the computer with an update or adaptation of its actions concerning the prediction of any event or controlling machine [14, 15]. Thus, enabling the implementation of ML in evaluating, predicting, and detecting fraud attempts in a banking system [16], it is possible to mitigate these adverse factor events, promoting greater credibility regarding the banking system and the organization [17–20].

Concerning the facts exposed, this study presents experiments using ML algorithms for a fraud classification prediction problem. For that, we use a database [21] that includes more than six million records monitored during one month of financial activities. In this context, an exploratory analysis of the data will be carried out in advance, leading to the treatment and implementation of a set of ML models to clarify the most favorable techniques aligned with the business models of the banking network.

The article is structured into four sessions. After the contextualization in section 1, section 2 provides a literature review approach focusing on the applications of ML related to fraud detection. Section 3 explores the case study, initially exposing a detailed exploratory analysis of the data, evidencing the main observations omitted in a large volume of data, and exposing the process of data processing for implementing the ML models. Section 4 concludes the study, exposing the gains of the approach and proposal for future works.

2. Machine Learning for Fraud Detection

Financial fraud in banking systems increasingly presents threats with high consequences in the corporate financial sector [22]. As stated in [4], in many cases, banking institutions are forced to improve their intelligence systems to detect fraudulent transactions continuously. In recent years, various studies have used machine learning and data mining techniques to enable alternative solutions to this scenario of high concern [23, 24].

As discussed in [25], in the last two decades, given the technological advance at the global level, there has been an immense growth in the field of artificial intelligence, especially regarding machine learning techniques, with improved access to the Internet, data and high-level computational processing [26, 27], favoring integration for training and real-time operations [28].

Generically, ML is a vast merging field in research and prediction and detection of fraud studies in computer systems [29, 30]. In this scenario, authors of empirical studies on fraud prediction have employed ML algorithms designed to improve the overall understanding of predictability models [31]. In addition, research on fraud classification models is integrated with neural networks, Bayesian networks, decision trees, and fuzzy logic, among other statistical techniques that classify the incidence of frauds and improve monetary processing systems [32, 33]. Figure 1 exposes an ML integration approach as support in detecting fraud in financial systems.

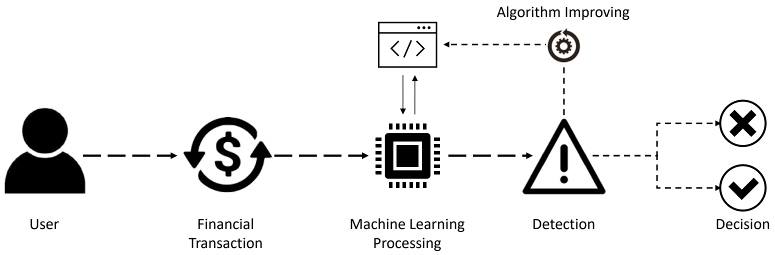

Fig. 1. Machine learning integration in financial systems

Within the scope of ML integration as support in detecting fraud in financial transactions, supervised learning and classification techniques are used as fraud analysis models, exposing misuse, transaction-level, and transactions categorized as fraud or standard transaction, based on learning a historical record set [34]. Examples of supervised learning techniques commonly applied to fraud detection are decision tree variants, neural networks, linear regression, logistic regression, KNN, and Naive Bayes [35].

The concept of using complex ML techniques mainly focuses on assisting laboratories or organizations in predicting and mitigating error as predictability analyses for a given scenario [36]. As stated in [37], predictive models and analysis, based on ML techniques, are typically used to predict future probabilities, indicating any unnatural pattern in the data or indicating risks at data points. In addition, all predictive analysis performs the integration of three fundamental components is:

- Database: the main component of the analysis is responsible for the quality and effectiveness of any predictive model;
- Mathematical Modeling: provides the treatment and construction of analyses enabling the manipulation of data and, in fact, the generation of predictive results through the construction of algorithms;
- Assumptions: High-value component that enables the integration of the database to the analyzed data, designing patterns according to historical events.

3. Case study

Regarding the case study, an exploratory analysis was implemented along with a framework related to machine learning in the scenario of evaluation and classification of different types of banking transactions regarding the presence of fraud. With a given framework, we seek to clarify, considering a database characteristic of Big Data, favorable ML models for computational integration in a banking system, enabling an operation based on artificial intelligence acting as an intervention model to a given fraud case before it occurs.

For the scenario in question, it was processed a database with more than 6 million bank transactions from a given international bank. It emphasizes that for a given study, the data in question are mischaracterized regarding customers' private information. The process of analysis and definition of the ML model establishes the most favorable among the tested within a methodological approach divided into four steps, being respective to the processes of exploratory analysis, data processing, and implementation of a set of ML models [38, 39]. For better clarification, Figure 2 exposes the methodological process in question.

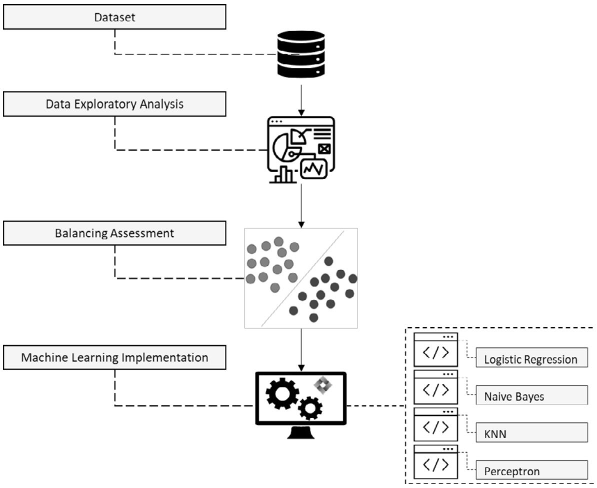

Fig. 2. Methodological process

In the database evaluated, eleven variables are established and recorded, they are:

- "step": indication of the period of monitoring of transactions in hour;
- "type": Type of bank transaction performed;
- "amount": Monetary value established in the bank transaction;
- "nameOrig": Code relating to the transaction source client;
- "oldbalanceOrig": Total monetary value present in the source account before the transaction;
- "newbalanceOrig": Total monetary value present in the source account after the transaction;
- "nameDest": Code relative to the target client of the transaction;
- "oldbalanceDest": Total monetary value present in the target account before the transaction;
- "newbalanceDest": Total monetary value present in the destination account after the transaction;
- "isFraud": Label of transactions defined as fraudulent;
- "isFlaggedFraud": Labels for the indication of fraud by the banking system before its implementation.

For the given evaluation process, the Python language was used to manipulate the data, along with the support of the models and statistical tools present in the *scikit-learn* library [40]. In this scenario, a descriptive data evaluation was obtained about the variables, as shown in Table 1.

Table 1. Descriptive analysis of the dataset

	Step	Amount	OldbalanceOrg	newbalanceOrig	oldbalanceDest	newbalanceDest	isFraud	isFlaggedFraud
mean	243.4	179861.9	833883.1	855113.7	1100701.7	1224996.4	0	0
Std	142.3	603858.2	2888242.7	2924048.5	3399180.1	3674128.9	0	0
Min	1.0	0	0	0	0	0	0	0
25%	156.0	13389.6	0	0	0	0	0	0
50%	239.0	74871.9	14208.0	0	132705.7	214661.4	0	0
75%	335.0	208721.5	107315.2	144258.4	943036.7	1111909.2	0	0
max	743.0	92445516.6	59585040.4	49585040.4	356015889.4	356179278.9	1	1

Regarding table 1, it notices that most transactions have high monetary value having an average value per transaction established at $ 179861.9, evidencing the need for integration of prediction models concerning the attempt of bank fraud. Moreover, the variable "isFlaggedFraud" only established 16 records.

In addition, it observes that among 6362620 transactions, 8213 cases defined as fraud, representing 0.13% of the total number of transactions, are frauds. In addition, it is also necessary to understand the monetary percentage frauded, thus leading to the loss of assets of the banking organization. In this context, the amount identified totalized more than 1 trillion dollars, and 1.05% of this value was lost in fraud, leading to a banking system of more than 12 billion dollars. Figure 3 brings a separate exposure to monetary distribution, where most regular transactions are between $0 and $250,000.00 and fraudulent transactions range from $150,000.00 to $1.5 million approximately.

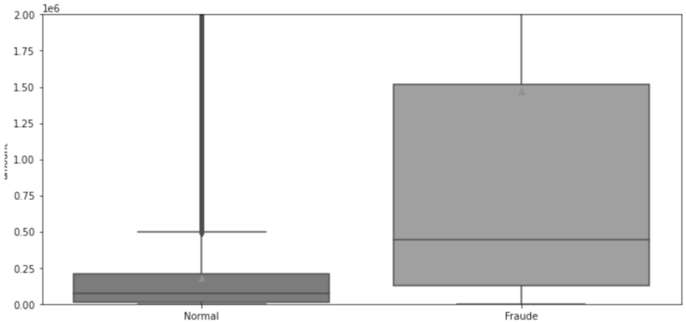

Fig. 3. The concentration of values concerning regular and fraud transactions

As for the types of bank transactions, are presented five models they are cash-out (35%), payment (34%), cash-in (22%), transfer (8.4%), and debit (0.6%). Even though there are multiple forms of banking transactions, the fraud records were established only cash out and transfer models, having the first 4116 and the second 4097 fraud records, approximately 50% of the cases each.

In addition to the analysis of processes, there were target accounts with more than one fraud record when the source accounts with a different account for each fraud record. Another point of interest relates to the total money withdrawal from the source account, where 98.1% of the transactions presented a given characteristic. Respective source accounts were observed that 65.2% were without any monetary value before the fraudulent transaction.

3.1. Pre-processing Unbalanced Data

Concerning unbalance in data between the records of fraud and normal transactions, there were used balancing techniques in addition to traditional machine learning models, allowing a closer analysis of the reality and a higher rate of effectiveness of the models [41, 42]. For the scenario in question, having 70% of the original base for training and 30% for tests, three balancing techniques were used, which were:

- Random under sampling (RUS): The model discards a random subset of the majority class, preserving the characteristics of the minority class, being favorable for large data volume [43];
- Synthetic Minority Oversampling Technique (SMOTE): it is an oversampling technique that, instead of simply replicating samples from the original minority set, generates synthetic samples based on the similarities between samples in the n-dimensional space of variables [43];
- Adaptive Synthetic (ADASYN): uses weighted distributions for different data samples of the minority class depending on how difficult it is for such samples to learn by models [43].

From the implementation of technical data, using the *python imblearn* library , it was possible to balance and build bases to be trained and evaluated in each scenario, with a total of four training bases, one unbalanced and three balanced by the techniques mentioned above, as shown in Figure 4.

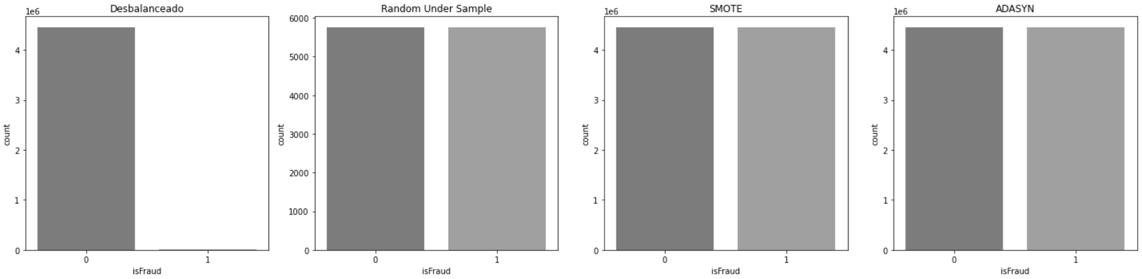

Fig. 4. Dataset balancing in three models

3.2. Machine Learning Analysis

After the balancing process, there were implemented four ML models with the aim of understanding/testing the effectiveness of the models and their integration feasibility for the fraud prediction problem. In this context, there were used the following models: Logistic regression, KNN, Naive Bayes, and Perceptron.

After training each of the ML models for each of the four bases, an understanding of the most favorable model for implementation is feasible, considering the analysis of the metrics of the area under roc curve (AUC), accuracy, number of false negatives, and number of false positives. Table 2 presents the results obtained within the above metrics in each evaluation scenario.

Table 2. Descriptive analysis from classification report

ML Model	Dataset Balancing	Accuracy	AUC	Number of False Positives	Number of False Negatives
Logistic Regression	Unbalanced	0.998	0.713	1877	1413
	RUS	0.908	0.908	174634	227
	SMOTE	0.912	0.910	168445	226
	ADASYN	0.836	0.913	312104	25
Naive Bayes	Unbalanced	0.992	0.584	13008	2031
	RUS	0.964	0.719	67465	1298
	SMOTE	0.963	0.721	68390	1288
	ADASYN	0.132	0.423	165922	701
KNN	Unbalanced	0.999	0.840	253	789
	RUS	0.941	0.952	12797	93
	SMOTE	0.995	0.945	8595	261
	ADASYN	0.995	0.945	8978	266
Perceptron	Unbalanced	0.992	0.883	15355	559
	RUS	0.687	0.838	597473	27
	SMOTE	0.650	0.819	668006	29
	ADASYN	0.866	0.919	256103	69

After implementing the ML models and obtaining their respective performance metrics, it was observed that there is no determining scenario as optimal in all metrics, thus exposing the presence of a trade-off n the evaluation.

With the evaluation of the accuracy and AUC metrics, the Logistic Regression and KNN models presented the best performances, especially about the training with balanced bases. However, when observing the performance of the metrics in the applications of the Naive Bayes algorithm, it is understood that a given model was not favorable in any of the cases, both for the unbalanced training base and the balanced ones.

A factor of importance in a given analysis is reflected in the classification of false negatives, representing the loss of money by the company, from the moment when the non-prediction of a fraudulent transaction will reflect in costs to the banking organization. The Logistic Regression algorithm presented the most favorable result for ADASYN balancing, preceded by the Perceptron algorithm, in the RUS and SMOTE training bases. However, it should be emphasized that even with low levels of false negatives, the applications presented for false positives, which in the case of banking transactions, may not mean direct monetary loss but can cause usability problems on the part of customers belonging to the banking network, considering that the negative classification of a normal transaction, would lead to interruption d the financial transaction.

In the presented scenario, it should be emphasized that a given model is not intended to be exhaustive but rather to present an approach to analyzing prediction algorithms regarding the integration of machine learning into a computer banking system, aiming at mitigating fraud. Finally, each ML model should be clarified as the most favorable, considering the business rules practiced by the organization and exposing the metrics most aligned with the organizational culture.

4. Conclusion

The article aimed to explore viable alternatives as a form of technological support in analyzing fraudulent transactions based on a series of the history of a banking organization. As a methodological approach, statistical analyses were used, exploring the data to clarify observations omitted in the database with more than six million financial transactions. The approach enabled a precise data analysis, providing the treatment and preparation for implementing predictive machine learning models of classification.

Using the Python language as a computational tool for implementing machine learning models, four different types of algorithms were implemented, applying each to four different training bases, one unbalanced and three balanced. With a given implementation, it was possible to obtain the metrics and understand the most favorable algorithms in terms of performances in this dataset, where the Logistic Regression and KNN models could be favorable to a future integration into the banking system. Even presenting good results, it is necessary to emphasize the need to evaluate the processing time for each algorithm, considering that the model should operate in real-time in the corporate system.

ML technologies are helpful regarding the feasibility of predictive analysis in bank fraud assessment scenarios. They can be useful for finding possible fraud attempts in a financial transaction system. For future studies, we suggest implementing other ML models and capturing other variables belonging to a banking transaction, thus providing greater effectiveness in the training of algorithms, and improving the insights and predictive capabilities of the models.

References

1. Costa, I.P. de A., Moreira, M.Â.L., Costa, A.P. de A., Teixeira, L.F.H. de S. de B., Gomes, C.F.S., Santos, M. Dos: Strategic Study for Managing the Portfolio of IT Courses Offered by a Corporate Training Company: An Approach in the Light of the ELECTRE-MOr Multicriteria Hybrid Method. International Journal of Information Technology & Decision Making. 1–29 (2021). https://doi.org/10.1142/S0219622021500565
2. Machkour, B., Abriane, A.: Industry 4.0 and its Implications for the Financial Sector. Procedia Computer Science. 177, 496–502 (2020)
3. Pereira, F. de C., Verocai, H.D., Cordeiro, V.R., Gomes, C.F.S., Costa, H.G.: Bibliometric Analysis of Information Systems Related to Innovation. Procedia Computer Science. 55, 298–307 (2015). https://doi.org/https://doi.org/10.1016/j.procs.2015.07.052
4. Santos, N., Rocha Junior, C. de S., Moreira, M.Â.L., Santos, M., Gomes, C.F.S., Costa, I.P. de A.: Strategy Analysis for project portfolio evaluation in a technology consulting company by the hybrid method THOR. Procedia Computer Science. 199, 134–141 (2022). https://doi.org/10.1016/j.procs.2022.01.017
5. Costa, I.P. de A., Costa, A.P. de A., Sanseverino, A.M., Gomes, C.F.S., Santos, M. dos: BIBLIOMETRIC STUDIES ON MULTI-CRITERIA DECISION ANALYSIS (MCDA) METHODS APPLIED IN MILITARY PROBLEMS. Pesquisa Operacional. 42, (2022). https://doi.org/10.1590/0101-7438.2022.042.00249414
6. Sadgali, I., Sael, N., Benabbou, F.: Performance of machine learning techniques in the detection of financial frauds. Procedia computer science. 148, 45–54 (2019)
7. Suh, J.B., Nicolaides, R., Trafford, R.: The effects of reducing opportunity and fraud risk factors on the occurrence of occupational fraud in financial institutions. International Journal of Law, Crime and Justice. 56, 79–88 (2019)
8. Nassim Mellem, P.M., de Araújo Costa, I.P., de Araújo Costa, A.P., Lellis Moreira, M.Â., Simões Gomes, C.F., dos Santos, M., de Pina Corriça, J.V.: Prospective scenarios applied in course portfolio management: An approach in light of the Momentum and ELECTRE-MOr methods. Procedia Computer Science. 199, 48–55 (2022). https://doi.org/10.1016/j.procs.2022.01.007
9. Maêda, S.M. do N., Basílio, M.P., Costa, I.P. de A., Moreira, M.Â.L., dos Santos, M., Gomes, C.F.S.: The SAPEVO-M-NC Method. Frontiers in Artificial Intelligence and Applications. 341, 89–95 (2021). https://doi.org/10.3233/faia210235
10. Jardim, R., dos Santos, M., Neto, E., Muradas, F.M., Santiago, B., Moreira, M.: Design of a framework of military defense system for governance of geoinformation. Procedia Computer Science. 199, 174–181 (2022). https://doi.org/10.1016/j.procs.2022.01.022
11. Jardim, R.R.-A.J., Santos, M., Neto, E.C. de O., da Silva, E.D., de Barros, F.C.M.M.: Integration of the waterfall model with ISO/IEC/IEEE 29148:2018 for the development of military defense system. IEEE Latin America Transactions. 18, 2096–2103 (2020). https://doi.org/10.1109/TLA.2020.9400437
12. Costa, I.P. de A., Basílio, M.P., Maêda, S.M. do N., Rodrigues, M.V.G., Moreira, M.Â.L., Gomes, C.F.S., dos Santos, M.: Algorithm Selection for Machine Learning Classification: An Application of the MELCHIOR Multicriteria Method. Frontiers in Artificial Intelligence and Applications. 341, 154–161 (2021). https://doi.org/10.3233/FAIA210243
13. Alqudah, N., Yaseen, Q.: Machine Learning for Traffic Analysis: A Review. Procedia Computer Science. 170, 911–916 (2020). https://doi.org/https://doi.org/10.1016/j.procs.2020.03.111
14. dos Santos, F.B., dos Santos, M.: Choice of armored vehicles on wheels for the Brazilian Marine Corps using PrOPPAGA. Procedia Computer Science. 199, 301–308 (2022). https://doi.org/10.1016/j.procs.2022.01.037
15. Moreira, M.Â.L., Gomes, C.F.S., Santos, M., Basilio, M.P., Costa, I.P. de A., Rocha Junior, C. de S., Jardim, R.R.-A.J.: Evaluation of drones for public security: a multicriteria approach by the PROMETHEE-SAPEVO-M1 systematic. Procedia Computer Science. 199, 125–133 (2022). https://doi.org/10.1016/j.procs.2022.01.016
16. Costa, I.P. de A., Basílio, M.P., Maêda, S.M. do N., Rodrigues, M.V.G., Moreira, M.Â.L., Gomes, C.F.S., Santos, M.: Bibliometric Studies on Multi-Criteria Decision Analysis (MCDA) Applied in Personnel Selection. Frontiers in Artificial Intelligence and Applications. 341, (2021). https://doi.org/10.3233/faia210239
17. Moreira, M.Â.L., Gomes, C.F.S., Santos, M., Silva Júnior, A.C., Costa, I.P. de A.: Sensitivity Analysis by the PROMETHEE-GAIA method: Algorithms evaluation for COVID-19 prediction. Procedia Computer Science. 199, 431–438 (2022). https://doi.org/10.1016/j.procs.2022.01.052

18. Tenorio, F.M., Santos, M. Dos, Gomes, C.F.S., Araujo, J.D.C., De Almeida, G.P.: THOR 2 Method: An Efficient Instrument in Situations Where There Is Uncertainty or Lack of Data. IEEE Access. 9, 161794–161805 (2021). https://doi.org/10.1109/ACCESS.2021.3132864
19. Moreira, M.Â.L., Gomes, C.F.S., Pereira, M.T., dos Santos, M.: SAPEVO-H2 a Multi-criteria Approach Based on Hierarchical Network: Analysis of Aircraft Systems for Brazilian Navy. Presented at the (2023)
20. Basílio, M.P., Pereira, V., Costa, H.G., Santos, M., Ghosh, A.: A Systematic Review of the Applications of Multi-Criteria Decision Aid Methods (1977–2022). Electronics. 11, 1720 (2022)
21. Roy, R.: Online Payments Fraud Detection Dataset, https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset, (2022)
22. Patil, S., Nemade, V., Soni, P.K.: Predictive Modelling For Credit Card Fraud Detection Using Data Analytics. Procedia Computer Science. 132, 385–395 (2018). https://doi.org/https://doi.org/10.1016/j.procs.2018.05.199
23. dos; Santos, M., Costa, I.P. de A., Gomes, C.F.S.: Multicriteria decision-making in the selection of warships: a new approach to the AHP method. International Journal of the Analytic Hierarchy Process. 13, (2021). https://doi.org/10.13033/ijahp.v13i1.833
24. Hilal, W., Gadsden, S.A., Yawney, J.: Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Systems with Applications. 193, 116429 (2022). https://doi.org/10.1016/j.eswa.2021.116429
25. Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., Lopez Garcia, A., Heredia, I., Malík, P., Hluchý, L.: Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. Artificial Intelligence Review. 52, 77–124 (2019)
26. Maêda, S.M., de Arajo Costa, I.P., Simões Gomes, C.F., dos Santos, M., da Mota, I.S., de Barros Teixeira, L.F.H. de S.: Economic and edaphoclimatic evaluation of Brazilian regions for African mahogany planting - an approach using the SAPEVO-M-NC ordinal method. Procedia Computer Science. 199, 323–330 (2022). https://doi.org/10.1016/j.procs.2022.01.196
27. Drumond, P., Basílio, M.P., Costa, I.P. de A., Pereira, D.A. de M., Gomes, C.F.S., dos Santos, M.: Multicriteria Analysis in Additive Manufacturing: An ELECTRE-MOr Based Approach. Presented at the October 29 (2021)
28. Rocha Junior, C. de S., Moreira, M.Â.L., Santos, M.: Selection of interns for startups: an approach based on the AHP-TOPSIS-2N method and the 3DM computational platform. Procedia Computer Science. 199, 984–991 (2022). https://doi.org/10.1016/j.procs.2022.01.124
29. de Oliveira, A.O., Oliveira, H.L.S., Gomes, C.F.S., Ribeiro, P.C.C.: Quantitative analysis of RFID' publications from 2006 to 2016. International Journal of Information Management. 48, 185–192 (2019).
30. Barros, M.D. de, Salles, C.A.L., Gomes, C.F.S., Silva, R.A. da, Costa, H.G.: Mapping of the Scientific Production on the ITIL Application Published in the National and International Literature. Procedia Computer Science. 55, 102–111 (2015). https://doi.org/https://doi.org/10.1016/j.procs.2015.07.013
31. Zhang, J., Pan, L., Han, Q.-L., Chen, C., Wen, S., Xiang, Y.: Deep learning based attack detection for cyber-physical system cybersecurity: A survey. IEEE/CAA Journal of Automatica Sinica. 9, 377–391 (2021)
32. Lokanan, M.E., Sharma, K.: Fraud prediction using machine learning: The case of investment advisors in Canada. Machine Learning with Applications. 8, 100269 (2022). https://doi.org/10.1016/j.mlwa.2022.100269
33. Maêda, S.M. do N., Basílio, M.P., Costa, I.P. de A., Moreira, M.Â.L., dos Santos, M., Gomes, C.F.S., de Almeida, I.D.P., Costa, A.P. de A.: Investments in Times of Pandemics: An Approach by the SAPEVO-M-NC Method. Presented at the October 29 (2021)
34. Srivastava, U., Gopalkrishnan, S.: Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks. Procedia Computer Science. 50, 643–652 (2015). https://doi.org/https://doi.org/10.1016/j.procs.2015.04.098
35. Singh, A., Jain, A.: An empirical study of aml approach for credit card fraud detection–financial transactions. International Journal of Computers Communications & Control. 14, 670–690 (2020)
36. Rocha Junior, C. de S., Moreira, M.Â.L., Santos, M., Gomes, C.F.S.: Creation and implementation of an IoT-based thermometer prototype for a food organization: case study. Procedia Computer Science. 199, 710–717 (2022). https://doi.org/10.1016/j.procs.2022.01.088
37. Singla, A., Jangir, H.: A comparative approach to predictive analytics with machine learning for fraud detection of realtime financial data. In: 2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3). pp. 1–4. IEEE (2020)
38. Gomes, C.F.S., Rodrigues, M.V.G., Costa, I.P. de A., dos Santos, M.: Ordering of Warships for the Brazilian Navy Using the New Method: AHP-Gaussian with Pearson's Correlation. Presented at the October 29 (2021)
39. Drumond, P., de Araújo Costa, I.P., Lellis Moreira, M.Â., dos Santos, M., Simões Gomes, C.F., do Nascimento Maêda, S.M.: Strategy study to prioritize marketing criteria: an approach in the light of the DEMATEL method. Procedia Computer Science. 199, 448–455 (2022). https://doi.org/10.1016/j.procs.2022.01.054
40. Bisong, E.: Introduction to Scikit-learn. In: Building machine learning and deep learning models on Google cloud platform. pp. 215–229. Springer (2019)
41. de Almeida, I.D.P., de Araújo Costa, I.P., de Araújo Costa, A.P., de Pina Corriça, J.V., Lellis Moreira, M.Â., Simões Gomes, C.F., dos Santos, M.: A multicriteria decision-making approach to classify military bases for the Brazilian Navy. Procedia Computer Science. 199, 79–86 (2022). https://doi.org/10.1016/j.procs.2022.01.198
42. Barbosa de Paula, N.O., de Araújo Costa, I.P., Drumond, P., Lellis Moreira, M.Â., Simões Gomes, C.F., dos Santos, M., do Nascimento Maêda, S.M.: Strategic support for the distribution of vaccines against Covid-19 to Brazilian remote areas: A multicriteria approach in the light of the ELECTRE-MOr method. Procedia Computer Science. 199, 40–47 (2022). https://doi.org/10.1016/j.procs.2022.01.006
43. Fernández, A., García, S., Galar, M., Prati, R.C., Krawczyk, B., Herrera, F.: Learning from imbalanced data sets. Springer (2018)