Revista

Cadernos de Finanças Públicas

03 | 2024

TESOURO NACIONAL

FORECASTING BID-RIGGING IN BRAZIL

Daniel Vitor Tartari Garruti

Universidade Federal de Uberlândia (UFU)

Vinícius Morais

Universidade Federal de Uberlândia (UFU)

Pablo Rogers

Universidade Federal de Uberlândia (UFU)

Flavio Barboza

Universidade Federal de Uberlândia (UFU)

ABSTRACT

Electronic bidding portals have modernized the acquisition of goods and services, making it more efficient, competitive and transparent. To keep up with these changes, oversight must also adapt. In the academic context, there has been progress in the use of Artificial Intelligence techniques, such as Random Forest, to predict bid-rigging. This study applied the model to 2022 tender data to predict the incidence of fines on companies. The model proved to be efficient, with an average monthly F1 Score of 78.5% and an annual F1 Score of 80%. Recall was 90% for both periods. In January, the best results were obtained (recall = 1.00 and F1 Score = 0.96), while in November there was a significant drop (recall = 0.47 and F1 Score = 0.54). The potential impact of this study not only on academia but also on citizens and public managers stands out, offering an effective tool for detecting potential misconduct.

Keywords: Tenders, Fraud, Public Service, Artificial Intelligence, Random Forest.
JEL: C53, D73, E66.

SUMMARY

1. INTRODUCTION 4

2. RELATED STUDIES 6

3. METHODOLOGY...9

3.1. Data collection 9

3.2. Random Forest 14

3.3. Algorithm...14

3.4. Analysis of results 15

4. RESULTS AND DISCUSSIONS 17

5. CONCLUSION...23

BIBLIOGRAPHICAL REFERENCES 26

1. INTRODUCTION

State economic activity plays a crucial role in boosting effective demand, stimulating production and job creation in the country. In Brazil, public sector procurement between 2006 and 2017, considering all federal entities, averaged 12.5% of GDP. Of this total, 6.8% corresponded to purchases by the Federal Government, of which Petrobras alone accounted for 4.2% (RIBEIRO; INÁCIO, 2019).

State contracts must be put out to tender (BRASIL, 1988). In this sense, the system of laws that regulates public procurement currently defines state purchasing power as an instrument capable of implementing public policies to reduce inequalities, promote socioeconomic development, innovation and sustainability (REJEB et al., 2023; ZAGO, 2018). Bidding is a formal procedure of the Public Administration, regulated by Law No. 14.133/2021, in which companies are called to submit proposals for the offer of goods and services, as established in a public notice (BRASIL, 2021).

Public procurement is one of the aspects of the state where corruption, fraud and collusion are most likely to occur (MAVIDIS; FOLINAS, 2022; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023; REJEB et al., 2023). Infringing conduct includes acts that compromise isonomy and competitiveness, such as combining prices or simulating bids; undue changes to the contract, payments that do not comply with the contractual rules, undue removal of a bidder through fraud or offering an advantage, unjustified delay in the execution of works or services, non-performance of the contract, corruption, obstruction of inspection. Transparency in contracting, the training of the team responsible and the work of internal and external control are key aspects in preventing fraud. In the literature, the main method for detecting fraud is the use of *Red Flags*/Alerts so that auditors can deepen their analysis of each case. Other forms of detection are associated with providing reporting channels and using multidisciplinary teams to investigate (GUNASEGARAN; BASIRUDDIN; RIZAL, 2023).

Brazilian regulations also establish a series of administrative and criminal sanctions to guarantee transparency, morality and efficiency in these processes. Law No. 8.666/1993, known as the Bidding and Contracts Law, provides for punishments such as a warning, fine, temporary suspension of participation in bidding and a declaration of ineligibility to bid or contract with the Public Administration (BRASIL, 1993). The Auction Law (Law No. 10.520/2002) and the Differentiated Public Procurement Regime (Law No. 12.462/2011) also stipulate similar sanctions, while the New Administrative Bids and Contracts Law (Law No. 14.133/2021) introduces new provisions, while maintaining a similar structure of administrative sanctions (BRASIL, 2002; BRASIL, 2011; BRASIL, 2021). The Anti-Corruption Law (Law No. 12.846/2013) establishes punishments for legal entities involved in harmful acts against the public administration, while the Criminal Code typifies crimes such as bid-rigging, embezzlement and corruption (BRASIL, 2013). The implementation of these laws aims to promote integrity in the corporate and governmental environment, as well as fighting corruption in the country, in line with international governance and transparency efforts.

However, the effective application of these laws faces significant challenges. The complexity of bidding processes and the large number of transactions make detecting irregularities an arduous task. Early identification of possible fraud and corrupt acts is essential for the effectiveness of preventive and punitive measures. In this context, technology plays a crucial role by providing advanced tools for monitoring and analyzing public procurement data.

The adoption of electronic systems has improved the efficiency and effectiveness of the public sector, especially in terms of transparency, governance, better use of public resources, promoting regional development and competition between suppliers (ADOBOR; YAWSON, 2023; GADOUR, 2024; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023; MAVIDIS; FOLINAS, 2022). Electronic procurement (*e-procurement*) has become commonplace in all spheres of government, with the Comprasnet portal standing out. Public administration is also increasingly adopting artificial intelligence (AI) tools. At the Federal Supreme Court, initiatives such as VitórIA and the Victor Project stand out, as well as RAFA 2030 (Artificial Networks Focused on the 2030 Agenda) (SUPREMO TRIBUNAL FEDERAL, 2021; SUPREMO TRIBUNAL FEDERAL, 2023a; SUPREMO TRIBUNAL FEDERAL, 2023b). The Federal Court of Auditors uses generative AIs, including the Zello chatbot, ChatTCU and Copilot TCU (TRIBUNAL DE CONTAS DA UNIÃO, 2024; SECOM TCU, 2020). In the field of public procurement, the main initiative is the Alice software (Bid, Contract and Tender Analyzer), developed by the Office of the Comptroller General (2024) (BIAGINI, 2024).

These electronic systems associated with Artificial Intelligence (AI) are fundamental tools for making procurement fraud more difficult (ADOBOR; YAWSON, 2023; GADOUR, 2024; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023). *Machine learning* (ML) models offer an innovative approach to predicting illegal conduct in public procurement, using analysis of large data sets to identify patterns and variables associated with corrupt behavior. These predictions can be used as risk scores by anti-corruption agencies and contracting parties to guide the selection of procurement operations that need stricter monitoring or audits.

In view of this, this study proposes a model for predicting fraud (alerts) in tenders in Brazil using the *Random Forest* (RF) technique. This study is useful for public agents working in public procurement, as it allows them to identify cases with a greater chance of irregularities in a predictive way, facilitating real-time analysis (SICILIANI et al., 2023). In view of the competencies performed by external control within the scope of public governance and the experience in the use of Artificial Intelligence (AI) tools, this work will be especially useful in audit work. From the perspective of academic contributions, this research fills a gap in the literature related to public procurement, as it seeks to explore the field of application of an AI tool, which is a field with few works, as identified by Rejeb et al. (2023) and Gunasegaran, Basiruddin and Rizal (2023).

2. RELATED STUDIES

Nai et al. (2023) applied ML techniques and recommender systems to improve the efficiency and transparency of public procurement processes, as well as to predict possible claims in administrative courts. The results of the study showed that it is possible to use ML techniques on real legal datasets related to procurement processes in Italy to predict claims in administrative courts based on the characteristics of procurement processes. The authors developed an ML model capable of recognizing anomalies in procurement processes and identifying the most relevant characteristics for classifying processes with or without claims. The model showed an accuracy of 80% in predicting claims, which demonstrates the effectiveness of the proposed approach.

In addition, the authors developed a recommendation system to return similar procurement processes and find companies for bidders, based on the requirements of the procurement process. The recommendation system was developed using ML techniques and deep neural networks, and showed promising results in identifying similar procurement processes and companies for bidders (NAI et al., 2023).

The results of the study highlight the applicability and potential of ML techniques and recommendation systems in improving procurement processes and predicting possible claims in administrative courts. These techniques can contribute to efficiency and transparency in the public sector, allowing authorities to identify possible anomalies in procurement processes and take preventative measures to avoid future claims. In addition, the study demonstrates the importance of exploiting open legal datasets to improve public sector systems, procedures and services (NAI et al., 2023).

Oliveira et al. (2023) identified signs of fraud by analyzing the interactions between participants in tenders, such as bidding companies and their partners, modeling these interactions as a social network. This methodology was successful in reducing a large volume of data into an indicator, which made it possible to rank the contracts where the risk of fraud was greatest. In this way, the audit work of the experts from the Minas Gerais Public Prosecutor's Office can be better oriented towards analyzing those cases where the risk of fraud is greatest.

Gallego et al. (2021) identified the demonstrated ability of ML models to predict illegal conduct in public procurement contracts, providing risk scores that can be used to prioritize contracts that require stricter monitoring. In addition, the study identified key variables, such as contract size and duration, implementation delays and sector standards, which are significantly associated with corruption and inefficiency.

As a result, the findings of this study offer interesting suggestions/proposals for the formulation of targeted policies and regulatory reforms to mitigate the risk of corruption in public procurement. The research also highlights the importance of transparency and accountability in the procurement process, proposing the use of web-based platforms to record and report public transactions. Furthermore, the combination of ML methods with traditional causal inference techniques is presented as a promising approach to strengthening anti-corruption policies and improving the integrity of public procurement processes.

Decarolis and Giorgiantonio (2020) emphasize the importance of improving data collection and transparency in public procurement as effective means of fighting corruption. They recommend adapting *red flag* indicators according to the specific characteristics of each region and sector, given the variability in the patterns of corruption observed. In addition, they highlight the potential of using advanced analytical tools, such as ML, in public administration to detect and prevent corruption more effectively.

Using ML models, they demonstrated that certain indicators, such as bidder uniqueness and lack of competition, are effective in predicting corruption. They also observed significant variations in these signals between different regions and types of contracts in Italy, suggesting the need for specific regional approaches to detecting corruption (DECAROLIS; GIORGIANTONIO, 2020).

Decarolis and Giorgiantonio (2020) indicate that in order to improve procurement processes and reduce corruption in public tenders, risk indicators must be identified, competition between companies must be improved and anti-corruption practices must be implemented. These measures can help strengthen integrity and transparency in public procurement, contributing to the reduction of corruption. In addition, a strong correlation has been identified between certain red flags and the increased risk of corruption in public procurement.

Henrique, Sobreiro and Kimura (2020) compared the main ML techniques, including RF, to classify Brazilian public contracts in relation to the risk of non-compliance. They also analyzed logistic regression, *K-Nearest Neighbours* (KNN), discriminant analysis and *Support Vector Machine* (SVM). Companies that received any serious sanctions were classified as high risk. The independent variables used were: procurement data, quantities of items, bids in electronic negotiations, average value approved in the bidding process, number of agencies with which contracts were concluded, number of federal entities with which agencies contracted with the company, number of winning items in the bidding process, disqualifications/irregularities suffered by the companies, average number of employees and their average salaries, income received by the partners, number of commercial activities registered with the SRF, age of the company, taxes paid to the authorities, size of the company (MESB and NPO), sum of the amounts negotiated in contracts in the favored condition, number of partners and employees who were also civil servants or in some kind of paid position, and any donations to political campaigns. The data was obtained from COMPRASNET and the Unified Supplier Registration System (SICAF), the Annual Social Information Report (RAIS) and the Superior Electoral Court (TSE). The model used 500 decision trees.

Similarly, Sá, Pessanha and Alves (2024) also compared the main classification algorithms, including RF, along with logistic regression, KNN, neural networks and SVM, to identify the propensity for fraud in the contracting processes of the Rio de Janeiro State Government. The dependent variable used was also the sanction suffered by the supplier. The independent variables were: CNPJ of the supplier, company size (Microenterprise or Small Business), quantity and value of contracts registered for the supplier, number of tenders in which the supplier participated, total tenders the supplier won, quantity of direct purchases registered for the supplier, value of direct purchases registered for the supplier, share capital, total number of CNAEs, distance between the supplier's headquarters and the place where the service is to be carried out, years since the company was set up, number of companies in which the supplier has partners in common, status with the Revenue Department (active or unfit/suspended), number of contracts signed with the supplier whose initial numbers failed the Benford test, average age of the company's partners. The data was taken from the Transparency Portal of the State of Rio de Janeiro and the Transparency Portal of the Federal Government. The model was configured by limiting the maximum depth of the tree to 8.

Both works by Henrique, Sobreiro and Kimura (2020) and Sá, Pessanha and Alves (2024) obtained good results for the RF, but did not calculate the main quality metrics of the classification model, such as precision, recall and F1 Score. Similarly, the academic papers by Faria (2023), Lopes (2019) and Silva (2022) also compared classification models, including RF, in the detection of bid-rigging in the Brazilian federal public administration. The results of Faria (2023) and Silva (2022) were in conflict with the results of Lopes (2019). The RF model had the best performance in the studies by Faria (2023) and Silva (2022), while it had the worst performance in the studies by Lopes (2019).

Recognizing the potential of the technique for Brazilian inspection bodies, this work seeks to evaluate the effectiveness of the Random Forest algorithm in detecting bid-rigging using data from the Transparency Portal of the Office of the Comptroller General.

3. METHODOLOGY

3.1. Data collection

To conduct this research, the data was obtained from the Transparency Portal, linked to the Office of the Comptroller General, during the period from 01/01/2022 to 31/12/2022. It was necessary to access four spreadsheets corresponding to each month, available in the Tenders and Contracts section, to obtain the information, as recommended by Gallego, Rivero and Martínez (2021). The data collected and the location on the website can be seen in Table 1.

Chart 1: Data collected from the Comptroller General's Transparency Portal.

INDEPENDENT VARIABLES	References
Budget	Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Type of entity	Aldana, Falcón-Cortés and Larralde (2022) Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Entity level	
Purpose of the tender	Gallego, Rivero and Martínez (2021)
Type of process	Aldana, Falcón-Cortés and Larralde (2022) Gallego, Rivero and Martínez (2021)
Type of **supplier**	Aldana, Falcón-Cortés and Larralde (2022) Sá, Pessanha and Alves (2024)
Day the contract was signed	Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Month the contract was signed	
Year the contract was signed	
Execution period	Aldana, Falcón-Cortés, Larralde (2022) Gallego, Rivero and Martínez (2021).
Waiting period	Gallego, Rivero and Martínez (2021)
Interval between date of signature and the date of granting	Gallego, Rivero and Martínez (2021)
Interval between start date of execution and the date of signature	Gallego, Rivero and Martínez (2021)

Source: Own elaboration

Once the aforementioned data had been obtained, the variables "Tender Number" and "UG Code" were combined, resulting in a new variable called "Code Number", in order to consolidate the data into a single table. This procedure was essential due to the identity of the bid number between various entities. Thus, by grouping the bid number with the entity code, we obtained a unique numbering for each process.

This resulted in 16 independent variables used in the research, which are listed in Table 2. In the literature, there is a wide range of variables used to study bid-rigging. However, most studies do not explicitly focus on explaining the reasons why companies become disqualified. Furthermore, the selection of independent variables is often not based on a strict logic of causality. Instead, many studies opt for exploratory or descriptive approaches, which aim to identify patterns of behavior without necessarily establishing direct causal relationships between the variables, as in the work of Henrique, Sobreiro and Kimura (2020).

In this sense, the work by Gallego, Rivero and Martínez (2021) stands out by proposing a set of basic independent variables, common to several other studies (see references Table 2), detailing an interpretation for the model and the importance of the variables. This set includes the variables (value, entities, companies, location, date, type of construction) pointed out by criminal experts as information that generally leads to a good estimate of the risk of fraud acquisition (LIMA et al., 2020). Thus, this study mainly used the set of predictor variables proposed by Gallego, Rivero and Martínez (2021), applied to the Brazilian context. The potential risk patterns associated with the variables are detailed below:

- **Budget:** Comparison between the planned budget and the final contract value can indicate overpricing. Figures that are much higher than expected may suggest collusion or corruption.
- **Type/Level of entity:** Government entities of different levels or types may have different patterns of fraud risk. The level of public and regulatory scrutiny can vary significantly between different government levels
- **Purpose of the bid:** The specific purpose may be associated with certain fraud risks.
- **Type of process:** Different processes (tender, competition, etc.) have different levels of transparency and associated risk. Less transparent processes can be more susceptible to fraud.
- **Type of supplier:** Suppliers with a history of irregularities or those that are newly established can increase the risk of fraud.
- **Day/Month/Year the contract was signed:** Dates close to the end of political terms or specific dates can be used to rush through hirings without due legal process.
- **Execution period:** Projects with long or short execution periods may indicate poor management or efforts to divert funds.
- **Waiting period:** Indicator of transparency and publicity signaling the diligence and punctuality of the authorities when publishing information on contracts. A very short waiting period between the announcement and the decision may indicate a lack of adequate competition
- **Interval between signature date and award date:** Very short intervals may suggest that the supplier was already predestined to win, regardless of competing bids.
- **Interval between execution start date and signature date:** Very short intervals may indicate insufficient preparation, while long intervals may be attempts to postpone execution without valid justification.
- **Distance to the nearest election:** Contracts signed too close to elections may be trying to influence election results or take advantage of the change of administration to go unnoticed.
- **Product category:** Certain product categories can be more prone to fraud, especially if they involve large sums of money or are of a more technical nature and difficult to assess.

• **State:** Regional variations in terms of governance and transparency can affect the risk of fraud. States with a history of corruption may have higher incidences of bid-rigging. In addition, the distance between the supplier's registered office and the place where the service is performed can indicate fraud.

Table 2: Variables organized by type, following the standard public management format. The variables highlighted in bold are those used in the final model.

INDEPENDENT VARIABLES	References
Budget	Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Type of entity	Aldana, Falcón-Cortés and Larralde (2022) Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Entity level	
Purpose of the tender	Gallego, Rivero and Martínez (2021)
Type of process	Aldana, Falcón-Cortés and Larralde (2022) Gallego, Rivero and Martínez (2021)
Type of **supplier**	Aldana, Falcón-Cortés and Larralde (2022) Sá, Pessanha and Alves (2024)
Day the contract was signed	Gallego, Rivero and Martínez (2021) Lima et al. (2020)
Month the contract was signed	
Year the contract was signed	
Execution period	Aldana, Falcón-Cortés, Larralde (2022) Gallego, Rivero and Martínez (2021).
Waiting period	Gallego, Rivero and Martínez (2021)
Interval between date of signature and the date of granting	Gallego, Rivero and Martínez (2021)
Interval between start date of execution and the date of signature	Gallego, Rivero and Martínez (2021)
Distance to closer election	Aldana, Falcón-Cortés and Larralde (2022) Gallego, Rivero and Martínez (2021) Henrique, Sobreiro and Kimura (2020) Velasco et al. (2021)
Product category	Gallego, Rivero and Martínez, (2021) Lima et al. (2020)
State	Abreu, Pereira and Gomes-Jr (2024) Gallego, Rivero and Martínez (2021) Sá; Pessanha and Alves (2024)
DEPENDENT VARIABLE	
Received a penalty registered with the comptroller general's office	Gallego, Rivero and Martínez (2021) Henrique, Sobreiro and Kimura (2020) Sá; Pessanha and Alves (2024)

Source: Own elaboration.

The dependent variable was defined as the institutions that received a penalty from the Comptroller General in 2022. It was chosen based on the work of Gallego, Rivero and Martínez (2021), Henrique, Sobreiro and Kimura (2020) and Sá, Pessanha and Alves (2024). Using the Comptroller General's penalty record as a dependent variable to indicate bid-rigging may not capture all existing cases of fraud. In fact, there may be situations of fraud that do not result in the application of a penalty due to various factors, such as limitations in investigations, lack of conclusive evidence or even delays in the inspection process. However, the choice of this dependent variable was based on the need for an objective and verifiable indicator of fraudulent behavior, as adopted in previous studies in the literature. In addition, the application of a penalty is a strong indication of significant and detected violations, which makes it a reliable metric for cases in which fraud is effectively identified and penalized. Further improvements to the model should incorporate other variables (such as ongoing investigations and complaints) that indicate irregularities, thus broadening the spectrum of fraud detection beyond penalties that have already been confirmed.

Table 3 shows the descriptive data for the predictor variables used in this study.

Table 3: Descriptive data for the variables.

VARIABLES	REMARKS	EXCLUSIVE VALUES	MOST COMMON VALUE	FREQUENCY
Bid amount	109196	52069	0	6195
Name UG	109196	2005	Federal University of Rio Grande do Sul	1510
Object	109196	98010	Information protected by secrecy under the terms of...	1012
Purchasing method	109196	10	Dispensation to Bid	59614
Purchase result date	109196	345	15/12/2022	886
UF	109196	28	RJ	22505
Winning name	950270	54509	J. J. VITALLI	7620
Date contract signed	35613	1056	01/09/2022	486
Effective date	35613	1111	01/09/2022	676
End date	35300	2162	31/12/2022	1194
DOU publication date	35613	270	26/12/2022	286
Description contract signature	35613	1056	01/09/2022	486
Purchase item description	4442683	10926	Building Maintenance/Refurbishment	72103

Fonte: Elaboração própria.

3.2. Random Forest

AI techniques have made significant progress in detecting bid-rigging (DECAROLIS; GIORGIANTONIO, 2020; GALLEGO et al., 2021; NAI et al., 2023; OLIVEIRA et al., 2023). Given this scenario, it was decided to use one of these techniques for forecasting in this study, specifically *Random Forest* (RF).

RF is made up of multiple decision trees, similar to a forest. The process of creating decision trees begins with the random generation of subsets of data from the originals, called *bootstrap* samples. Randomization is then used again to select the predictor variables (attributes), forming the decision nodes. The nodes are defined based on the characteristics that best separate the data according to the desired result, establishing cut-off points in the independent variables in order to maximize differentiation between the categories. Metrics such as Gini impurity, information gain or mean square error (MSE) can be used to assess the quality of the split. The recursive separation of the data forms (mostly) pure nodes containing data from only one class, called leaf nodes. This procedure is repeated to form several decision trees, each with different feature structures, resulting in different models. Finally, to obtain a final prediction, the algorithm goes through each tree and determines the result of each variable, selecting the most frequent result through a "voting" process.

The model's developer, Breiman (2001) highlights the advantages of the RF technique, emphasizing its accuracy, ability to deal with *outliers* and noise, and ease of implementation. He points out that the technique provides adequate estimates for internal variables, such as error, strength, correlation and importance of variables, while avoiding problems of *overfitting* (failure to generalize), ensuring reliability in predicting future results.

3.3 Algorithm

The RF model was implemented in Python 3, using the *RandomForestClassifier* class from the *sklearn.ensemble* module of the *Scikit-Learn* library (SCIKIT-LEARN, 2024). After checking the distribution of the classes, the data was balanced to balance the under-represented and over-represented classes, reducing the number of examples in the majority class. The data was divided into sets of 70% for training and 30% for testing. The model was set up with 100 trees (n_estimators=100) and random_state=42. The accuracy of the model was assessed using the confusion matrix and the classification report. Stratified cross-validation was also carried out to check the model's effectiveness in terms of its ability to generalize to new data, dividing the data set into 5 parts (*n_splits* = 5). Finally, it was analyzed which characteristics are most important for predicting fraud and the ROC curve was generated.

Analysis was initially carried out on a monthly basis, due to computational constraints (2-core processor and 4GB RAM) to process the volume of data. Later, with a more powerful computer (12-core processor and 32GB RAM) it was possible to expand the analysis to the annual interval, allowing for a more comprehensive evaluation.

3.4 Analysis of results

The performance of the classification model implemented was evaluated by analyzing type I and II errors, accuracy, precision, sensitivity (*recall*), *F1 Score*, as well as comparing it with the results of the studies mentioned in the theoretical framework.

Table 4 below shows the count of events and predictions so that you can see the quality of the predictive model, also known as the confusion matrix. The confusion matrix makes it possible to clearly visualize the different types of errors (type I and type II) and the correct results of the model. This makes it easier to analyze the model's accuracy in predicting fraud, as well as to identify processes that meet the assumptions of legality, i.e. that there is no evidence of fraud.

Table 4: Confusion Matrix, which will show the results of the predictions and the count of hits and misses that are used to assess predictive quality.

Actual versus forecast	**Predicted as Fraud**	**Forecast as Non-Fraud**
Fraud occurred	True Positive (TP)	False Negative (FN) *Type II error*
Fraud Did not occur	False Positive (FP) *Type I error*	True Negative (TN)

Source: Own elaboration.

More specifically, type I error (given by Equation 1) occurs when fraud has not been detected in the bidding process, but is wrongly pointed out by the forecasting model. Erring in this case is considered less costly, since it is like a warning sign for more careful verification and is therefore an integral (and almost natural) part of the activities of inspection bodies. Type II error (Equation 2) occurs when fraud is detected but not predicted by the model.

$$Erro\ do\ Tipo\ I = \frac{FP}{FP+TN} \quad (1)$$

$$Erro\ do\ Tipo\ II = \frac{FN}{FN+TP} \quad (2)$$

Accuracy (Equation 3) represents the model's overall average success rate. However, it is an insufficient measure, especially for unbalanced classes. A more in-depth analysis of the model's effectiveness was carried out by also analyzing accuracy and sensitivity.

$$Acurácia = \frac{TN+TP}{TP+FP+TN+FN} \quad (3)$$

Accuracy and sensitivities (Equations 4) are complementary metrics. The choice of which metric to emphasize depends on the specific context of the application and the desired balance between detecting positives and avoiding false positives. In bid-rigging detection, sensitivity (*recall*) is often considered more critical. False negatives, i.e. frauds that go undetected, are associated with financial losses, reputational damage and foster corruption networks. Prioritizing sensitivity helps ensure that the majority of frauds are detected, minimizing the chance of frauds going undetected. Although less critical, accuracy is still important, especially in contexts where the volume of transactions is high. High false positive rates can lead to operational overload, increasing the costs of unnecessary investigations and potentially damaging the experience of companies/suppliers if legitimate transactions are frequently blocked or delayed. Increasing accuracy reduces the number of false positives, which can help maintain operational efficiency.

Thus, the F1 Score is particularly valuable because it calculates the harmonic mean between precision and sensitivity (*recall*), allowing the identification of the model that offers the best value for money. The F1 score ranges from 0 to 1, where 1 is the best possible value and 0 is the worst. A value of 1 indicates perfect classification, with perfect accuracy and sensitivity, meaning that all true positives and no false positives have been identified. A value of 0 indicates that the model has completely failed to identify true positives, or that all positive results are false positives.

$$Precisão = \frac{TP}{TP+FP} \qquad Sensibilidade = \frac{TP}{TP+FN} \quad (4)$$

$$F1\_Score = 2.\frac{Precisão \times Sensibilidade}{Precisão + Sensibilidade} \quad (5)$$

In addition, the RF technique has an interesting property that allows us to examine which predictor variables were most important in making the predictions. Thus, this was also a part of the analysis that could be considerably useful for its users to pay more attention to certain signposts and thus increase their efficiency and improve the inspection process in bidding processes.

The *Receiver Operating Characteristic* (ROC) curve makes it possible to graphically evaluate the classifier's performance, relating sensitivity and specificity. The efficiency of the classifier is observed as the curve approaches the ideal classifier: false positive rate equal to zero and true positive rate equal to 1, corresponding to the upper left point of the graph, coordinates (0.1) and the *area under* the ROC *curve* (AUC) ideally close to 1. Thus, the classifier is effective if AUC >0.5 and the ROC curve is above the diagonal dashed line, which represents a random classifier.

4. RESULTS AND DISCUSSIONS

The initial tests were carried out using the 16 independent variables listed in Table 2. However, in order to improve computational performance, the independent variables of least relevance to the model were excluded: type of entity, level of entity, type of supplier, month the contract was signed and year the contract was signed. The final model used 11 variables to predict fines, indicating fraud.

The results of the RF model for the set of variables selected, considering the year 2022, are shown in Tables 1 to 4. The data was analyzed monthly and annually using the confusion matrix (Tables 1 and 2) and the measures of accuracy, precision, *recall* and *F1 Score* (Tables 3 and 4).

The confusion matrices obtained using the model are shown in Table 1 e Table 2. It can be seen that for the annual analysis, type I error (when the model identifies a false positive) and type II error (when the model fails to identify a negative) do not significantly affect the model, considering that the accuracy was 0.78 and the *F1 Score* was greater than 0.74. With regard to the monthly matrices, the overall average accuracy and *F1 Score* were higher than 0.7. When analyzing each month in detail, we see situations such as July. In this case, the metrics indicate that the model has good sensitivity (*recall*), correctly identifying the majority of positive cases, but precision is lower (0.66), suggesting that there are a significant number of false positives. The overall accuracy is reasonable, but there may be room for improvement, especially in relation to precision.

TABLE 1: Monthly confusion matrix (10% sample)

January	Positive prediction	Negative prediction	**February**	Positive prediction	Negative prediction	**March**	Positive prediction	Negative prediction
Real positive	590	54	Real positive	2927	2702	Real positive	4603	2849
Real negative	0	595	Negative real	548	4964	Negative real	698	6752
April	Positive prediction	Negative prediction	**May**	Positive prediction	Negative prediction	**June**	Positive prediction	Negative prediction
Real positive	2202	458	Real positive	17155	6785	Real positive	4854	1124
Real negative	17	2658	Real negative	4462	19609	Real negative	76	5927
July	Positive prediction	Negative prediction	**August**	Positive prediction	Negative prediction	**September**	Positive prediction	Negative prediction
Real positive	1694	1665	Real positive	535	149	Real positive	1851	1404
Real negative	119	3221	Real negative	17	618	Real negative	665	2565
October	Positive prediction	Negative prediction	**November**	Positive prediction	Negative prediction	**December**	Positive prediction	Negative prediction
Real positive	14357	7035	Real positive	28604	9667	Real positive	4457	849
Real negative	4391	16933	Real negative	20546	17922	Real negative	523	4742

Source: Own elaboration.

TABLE 2: Annual confusion matrix (10% sample)

2022	Positive prediction	Negative prediction
Real positive	114501	60487
Real negative	18011	156264

Source: Own elaboration.

A Table 4 shows the results of the model for a 10% sample of the data, considering the information for the entire period, i.e. the year 2022. The average accuracy was 0.78, slightly higher than the monthly average (see Table 3). The model's *F1 Score* was 0.74 for undetected irregularities and 0.8 for detected irregularities.

The information in Table 3 was generated with a 10% sample of the data. The average accuracy was 0.76, with the highest value being 0.96 in January and the lowest value being 0.61 in November. This means that the model was correct 76% of the time.

TABLE 3: RF model results by month (10% sample)

Month	Situation	Precision	*Recall*	F1-Score	Support
January	Irregularity not detected (0)	1	0,92	0,96	644
	Irregularity detected (1)	0,92	1	0,96	595
	Accuracy			0,96	1239
February	Irregularity not detected (0)	0,84	0,52	0,64	5629
	Irregularity detected (1)	0,65	0,9	0,75	5512
	Accuracy			0,71	11141
March	Irregularity not detected (0)	0,87	0,62	0,72	7452
	Irregularity detected (1)	0,7	0,91	0,79	7450
	Accuracy			0,76	14902
April	Irregularity not detected (0)	0,99	0,83	0,9	2660
	Irregularity detected (1)	0,85	0,99	0,92	2675
	Accuracy			0,91	5335
May	Irregularity not detected (0)	0,79	0,72	0,75	23940
	Irregularity detected (1)	0,74	0,81	0,78	24071
	Accuracy			0,77	48011
June	Irregularity not detected (0)	0,98	0,81	0,89	5978
	Irregularity detected (1)	0,84	0,99	0,91	6003
	Accuracy			0,9	11981
July	Irregularity not detected (0)	0,93	0,5	0,66	3359
	Irregularity detected (1)	0,66	0,96	0,78	3340
	Accuracy			0,73	6699
August	Irregularity not detected (0)	0,97	0,78	0,87	684
	Irregularity detected (1)	0,81	0,97	0,88	635
	Accuracy			0,87	1319
September	Irregularity not detected (0)	0,74	0,57	0,64	3255
	Irregularity detected (1)	0,65	0,79	0,71	3230
	Accuracy			0,68	6485
October	Irregularity not detected (0)	0,77	0,67	0,72	21392
	Irregularity detected (1)	0,71	0,79	0,75	21324
	Accuracy			0,73	42716
November	Irregularity not detected (0)	0,58	0,75	0,65	38271
	Irregularity detected (1)	0,65	0,47	0,54	38468
	Accuracy			0,61	76739
December	Irregularity not detected (0)	0,89	0,84	0,87	5306
	Irregularity detected (1)	0,85	0,9	0,87	5265
	Accuracy			0,87	10571
Average	Irregularity not detected (0)	0,88	0,735	0,735	94630
	Irregularity detected (1)	0,725	0,905	0,785	94497
	Accuracy			0,765	189127

Source: Own elaboration

TABLE 4: RF model results (10% sample)

Year	Situation	Precision	*Recall*	F1-Score	Support
2022	Irregularity not detected (0)	0,86	0,65	0,74	174988
	Irregularity detected (1)	0,72	0,9	0,8	174275
	Accuracy			0,78	349263

Source: Own elaboration

The accuracy values were similar to those obtained by Henrique, Sobreiro and Kimura (2020) at 75.58% and lower than those obtained by Sá, Pessanha and Alves (2024) at 95.83%.

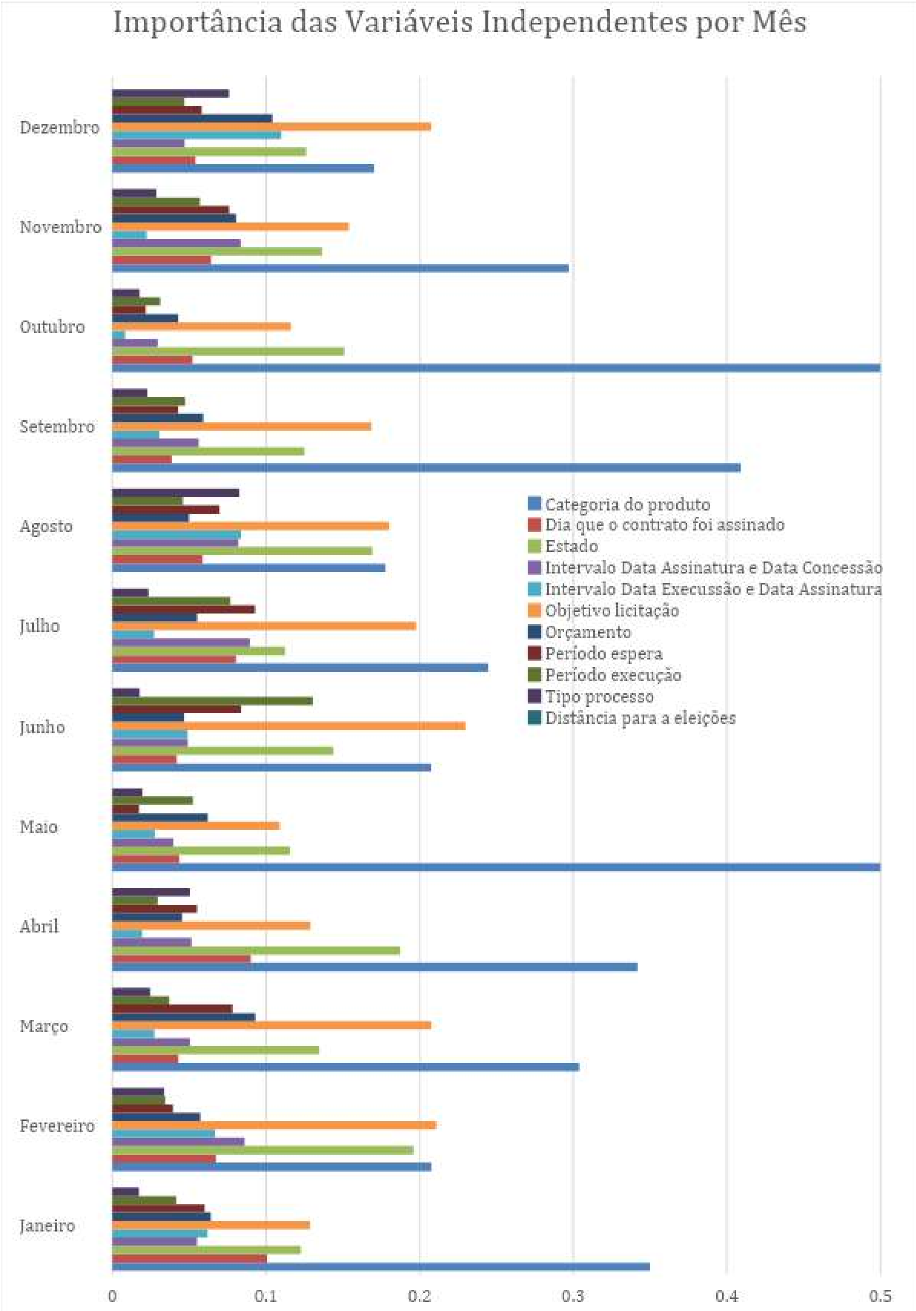
In addition, with regard to the *F1 Score*, the model showed average values of 0.72 for undetected irregularity and 0.79 for detected irregularity. These values indicate that the model is making good predictions for both the positive and negative classes. For the negative class, the model has significant difficulties in correctly identifying cases that are not irregular, as indicated by the low *recall*. This suggests that many legitimate cases may be wrongly flagged as irregular, potentially causing unnecessary disruptions and investigations. For positive class, the model was effective in detecting irregularities, with a high *recall*, which is positive for fraud detection. However, the lower precision implies a considerable number of false positives, which can also lead to additional operating costs and possible inconvenience.

The lowest value obtained for *F1 Score* was 0.54 for the irregularity detected in November. It should be noted that this was the only value lower than 0.6 for this coefficient. In November, there was a significant drop in precision for non-irregular cases and an increase in *recall* for the same class, indicating that the model began to identify legitimate cases more effectively, but with a corresponding increase in false positives. For cases of irregularities, there was a considerable reduction in *recall* (from 1.00 to 0.47), which means that the model no longer identifies many real cases of irregularities. In fact, there was a change in the model's pattern from September to November compared to previous months.

The last quarter of the year is traditionally a period of intense public bidding activity in Brazil. Many agencies seek to use remaining budgets to ensure that funds are not wasted or reduced in the next fiscal cycle. In 2022, this may have been further intensified by the presidential elections, which were marked by strong political polarization. The outcome of the elections at the end of October and the change of government may have influenced a change in the pattern of bidding data, reflected in a worsening of the forecast model observed especially in the month of November. Another possible explanation that may have influenced the model's poor performance in this month was the considerably higher volume of data used, corresponding to 32% of the total data, which may have caused the model to be under-adjusted.

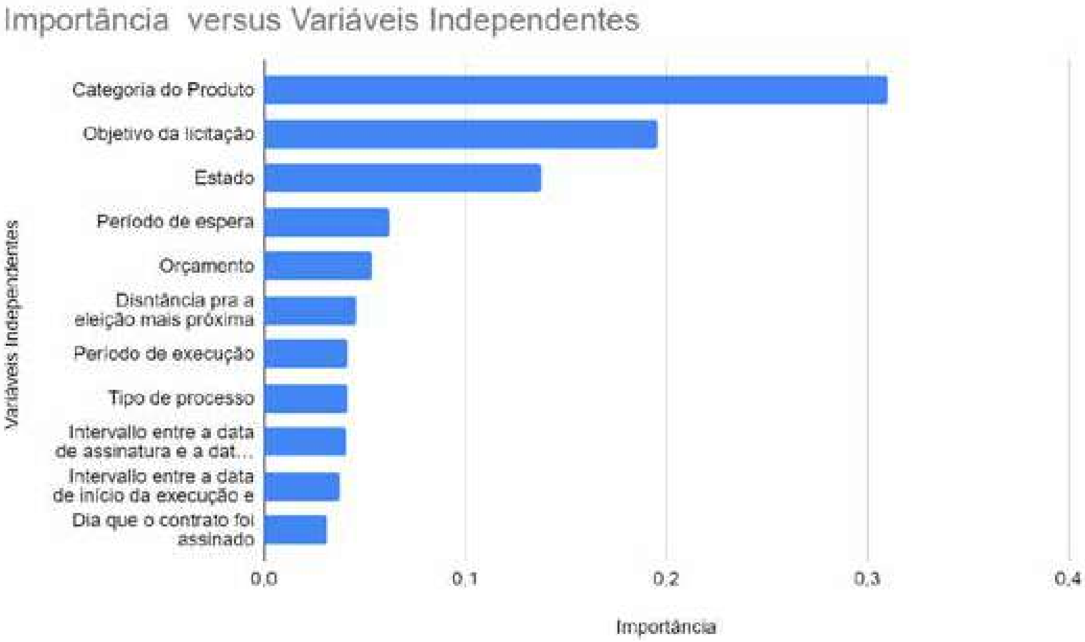
Another result obtained from the RF model is a list of the importance of the independent variables (Figure 1 and Figure 2). For the annual analysis, Product Category, Bidding Objective and State are the variables that contribute most to the model's accuracy, accounting for 64.40% of the total. Excluding the three variables with the greatest contribution to the model, there is a certain balance between the importance of the other variables, with the best result being 0.06 and the worst result being 0.03. Of this second group, the most important variables are Budget (0.06), Waiting Period (0.05) and Distance to the nearest election (0.05).

FIGURE 1: Importance of independent variables by month (10% sample)



Source: Own elaboration.

Figure 2: Importance of annual independent variables (10% sample)



Source: Own elaboration.

Although this study used variables from the work of Gallego, Rivero and Martínez (2021), they used different classification models, comparing lasso and GBM, with data from Colombia. They obtained as key independent variables the typical characteristics of contracts such as their size (measured by estimated budget and duration), delays in implementation, time before the next election, and geographical and sector-specific patterns. Although overall the importance of the variables was similar between the two studies, in Gallego, Rivero and Martínez (2021) Budget was the most relevant variable, while this study found Product Category to be the most significant. This may indicate that different geographical contexts with specific economic, political and cultural factors can affect the relevance of the variables. Or the difference may simply be related to differences in the models used, something that should be looked at in depth when choosing the method, suggesting that certain classification models may be better at detecting certain contexts.

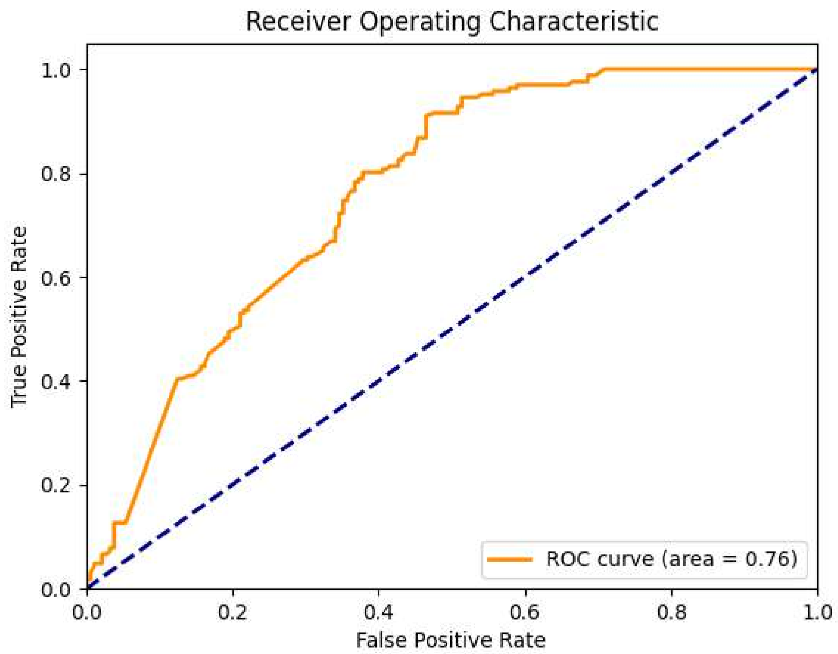
Sá, Pessanha and Alves (2024) using RF with data from Rio de Janeiro, obtained similar results to Gallego, Rivero and Martínez (2021), where the main variable was the 'Value of contracts registered for the supplier in the period', followed by 'Quantity of contracts registered for the supplier in the period' and share capital.

The monthly analysis (Figure 1) showed variations in the importance of the independent variables throughout the year. For example, the Budget variable, which came fourth in importance in the annual analysis, with the exception of August and December, was less important in all the other months.

In the monthly analysis, the variable 'Distance to the nearest election' was not relevant in any of the months studied. This is due to the fact that, within each month, the distance to the next election is constant for all the data analyzed, eliminating variations that could highlight its importance in the model.

The *area under the curve* (AUC) achieved for a test sample with 0.01% of the annual data was 76%, slightly higher than that obtained by Henrique, Sobreiro and Kimura (2020) (74.92%). The ROC curve obtained is shown in Figure 3.

FIGURE 3: ROC curve for 0.01% sample of annual data



Source: Own elaboration.

5. CONCLUSION

The computerization of bidding processes has brought greater transparency and efficiency, reducing operating costs and speeding up the contracting process. In addition, it has promoted broader competition. However, this progress has also made it necessary for auditing bodies to adopt new tools for detecting fraud, adapting to technological innovations in the sector and the massive volume of data involved. Thus, there is a growing trend towards the application of AI and ML techniques in the inspection of electronic public procurement systems.

Classification models make it possible to predict the occurrence of fraud, based on pre-

-identified patterns in a labeled data set, and to point out the main variables indicating irregularities, directing investigation efforts. Among these techniques, the characteristics of the RF model are particularly suitable for this application, due to its robustness, accuracy, ability to handle the complexity of the data (large volume, high dimensionality and missing values), as well as its simplicity of implementation. It also stands out for providing a measure of the importance of variables. Several comparative studies between the main algorithms have proven its efficiency in the context of detecting fraud in Brazilian tenders (HENRIQUE; SOBREIRO; KIMURA, 2020; SÁ; PESSANHA; ALVES, 2024) and others have indicated that it is the best classification model (FARIA, 2023; SILVA, 2022).

The RF model was implemented using data from the Transparency Portal of the Office of the Comptroller General, from 01/01/2022 to 31/12/2022. A set of 11 independent variables was used: budget, purpose of the tender, type of process, day the contract was signed, execution period, waiting period, interval between signature date and award date, interval between execution start date and signature date, distance to the nearest election, product category, state. The dependent variable was defined as fines received by the companies. The data was divided into 70% for training and 30% for testing. The classifier was implemented with 100 trees.

The implemented classifier performed efficiently with an average monthly *F1 Score* of 78.5% and an annual *score* of 80%. It stood out for its ability to detect the majority of real cases of fraud, with an average monthly and annual *recall* of 90%. The most important variables for the model were Product Category, Bidding Objective and State for both the monthly and annual analyses. However, the importance of these variables fluctuated significantly over the months.

There was a significant change in pattern between the months, which should be investigated further. For cases of irregularities, there was a considerable reduction in the quality parameters with maximum values in January and minimum values in November (change in *recall* from 1.00 to 0.47, respectively, and change in *F1 Score* from 0.96 to 0.54, respectively). Although the variable 'distance to the nearest election' is only sixth in importance, the 2022 presidential elections are believed to have had a significant influence on the change in pattern.

The research was limited by computational constraints. The large volume of data significantly increased processing time, requiring the data to be divided up by month in order to carry out the analysis. Annual processing required a computer with at least 32GB of RAM, and even then it took approximately 1 hour to complete the data processing. Furthermore, the dependent variable of fines received by companies alone does not capture all existing cases of fraud. On-going investigations or complaints are possible dependent variables to be incorporated into the model in future work to broaden the spectrum of fraud detection. The importance of other independent variables should also be assessed. Finally, future analyses should include other years to compare performance and refine the model.

The RF model developed has the potential for practical application beyond academia, offering public managers an effective tool for detecting possible misconduct in the bidding process.

BIBLIOGRAPHICAL REFERENCES

ABREU, B. M.; PEREIRA, T. H. S.; GOMES-JR, L. Fraud Detection in Public Tenders: A Comparison of Anomaly Detection Models. In: REGIONAL DATA BANKING SCHOOL (ERBD), 2024, Farroupilha/RS. Proceedings [...]. Porto Alegre: **Brazilian Computer Society**, p. 81-90, 2024. DOI: https://doi.org/10.5753/erbd.2024.238821.

ADOBOR, H.; YAWSON, R. The promise of artificial intelligence in combating public corruption in the emerging economies: A conceptual framework. **Science and Public Policy**, v. 50, n. 3, p. 355-370, 2023. DOI: https://doi.org/10.1093/scipol/scac068.

ALDANA, A.; FALCÓN-CORTÉS, A.; LARRALDE, H. A machine learning model to identify corruption in México's public procurement contracts. arXiv preprint arXiv:2211.01478, 2022.

AMORIM, V. A. J. Tenders and administrative contracts: theory and case law. Brasília: Federal Senate, 2020. Available at: https://bibliotecadigital.stf.jus.br/xmlui/bitstream/handle/123456789/2461/1182434.pdf?sequence=1&isAllowed=y. Accessed on: July 24, 2024.

BIAGINI, E. AI tool has already avoided spending R$11 billion on suspicious bids, says CGU minister. **Agência Gov**, 2024. Available at: <https://agenciagov.ebc.com.br/noticias/ferramenta-de-ia-da-cgu-ja-economizou-r-11-bilhoes-em-gastos-publicos-em-licitacoes-suspeitas-diz-ministro>. Accessed on: August 2, 2024.

BRAZIL. Constitution (1988). Constitution of the Federative Republic of Brazil of 1988. Promulgated on October 5, 1988. Brasília, DF: Federal Senate, Graphic Center, 1988.

BRAZIL. Law No. 8.666, of June 21, 1993. Establishes rules for public administration bids and contracts and makes other provisions. Federal Official Gazette, Brasília, DF, June 22, 1993. Available at: https://www.planalto.gov.br/ccivil_03/leis/l8666cons.htm.

BRAZIL. Law No. 10.520/2002, of July 17, 2002. Establishes, within the scope of the Union, States, Federal District and Municipalities, under the terms of art. 37, item XXI, of the Federal Constitution, a bidding method called auction, for the acquisition of common goods and services, and makes other provisions. Diário Oficial da União, Brasília, DF, July 17, 2002. Available at: https://www.planalto.gov.br/ccivil_03/LEIS/2002/L10520.htm?origin=instituicao.

BRAZIL. Law No. 12.462, of August 4, 2011. Establishes the Differentiated Public Procurement Regime - RDC; amends Law No. 10.683, of May 28, 2003, which organizes the Presidency of the Republic and the Ministries, and makes other provisions. Diário Oficial da União, Brasília, DF, Aug. 5, 2011. Available at: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12462.htm.

BRAZIL. Law No. 12.846, of August 1, 2013. Provides for the administrative and civil liability of legal entities for the practice of acts against the public administration, national or foreign, and makes other provisions. Federal Official Gazette, Brasília, DF, Aug. 2, 2013. Available at: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12846.htm.

BRAZIL. Law No. 14.133, of April 1, 2021. Bidding and Administrative Contracts Law. Federal Official Gazette, Brasília, DF, April 1, 2021. Available at: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14133.htm.

BREIMAN, Leo. Random forests. **Machine learning**, v. 45, p. 5-32, 2001. https://link.springer.com/article/10.1023/A:1010933404324

COMPTROLLER GENERAL OF THE UNION. Alice - Bid, Contract and Public Notice Analyzer. Available at: https://www.gov.br/cgu/pt-br/assuntos/auditoria-e-fiscalizacao/alice. Accessed on: August 2, 2024.

DECAROLIS, F.; GIORGIANTONIO,. Corruption red flags in public procurement: new evidence from Italian calls for tenders. **EPJ Data Science**, v. 11, n. 1, p. 16, 2022. https://doi.org/10.1140/epjds/s13688-022-00325-x

FARIA, G. A. Fraud in Government Procurement: Detection with Machine Learning. 2023. Course Conclusion Paper (Graduation in Public Administration) - School of Public Administration, Center for Legal and Political Sciences, Federal University of the State of Rio de Janeiro, Rio de Janeiro, 2023.

GADOUR, M. M. E. Corruption in public procurement: Can e-procurement and artificial intelligence make a difference in Africa? **QScience Connect**, v. 2024, n. 1-Thesis, p. 2, 2024. DOI: https://doi.org/10.5339/connect.2024.spt.2.

GALLEGO, J.; RIVERO, G.; MARTÍNEZ, J. Preventing rather than punishing: An early warning model of malfeasance in public procurement. **International Journal of Forecasting**, v. 37, n. 1, p. 360-377, 2021. http://dx.doi.org/10.1016/j.ijforecast.2020.06.006.

GUNASEGARAN, M.; BASIRUDDIN, R.; RIZAL, A. M. Detecting and Preventing Fraud in e-procurement of Public Sector: A Review, Synthesis and Opportunities for Future Research. **International Journal of Academic Research in Business and Social Sciences**, v. 13, n.1, p. 1444 - 1463, 2023.

HENRIQUE, B. M.; SOBREIRO, V. A.; KIMURA, H. Contracting in Brazilian public administration: A machine learning approach. **Expert Systems**. 2020. https://doi.org/10.1111/exsy.12550

LIMA, M. C.; SILVA, R. C.; MENDES, F. L. S.; CARVALHO, L. R.; ARAUJO, A.; VIDAL, F. B. Inferring about fraudulent collusion risk on Brazilian public works contracts in official texts using a Bi-LSTM approach. In: COHN, T.; HE, Y.; LIU, Y. (Ed.). Findings of the Association for Computational Linguistics: EMNLP 2020. Online: **Association for Computational Linguistics**, p. 1580-1588, 2020. Available at: https://aclanthology.org/2020.findings-emnlp.143. Accessed on: July 22, 2024. DOI: 10.18653/v1/2020.findings-emnlp.143.

LOPES, M. A. Application of machine learning in the detection of public fraud. 2019. Dissertation (Master of Business Administration) - School of Economics, Business Administration and Accounting, University of São Paulo, São Paulo, 2019. doi:10.11606/D.12.2020.tde-10022020-174317. Accessed on: 2024-07-22.

MAVIDIS A, FOLINAS D. From Public E-Procurement 3.0 to E-Procurement 4.0; A Critical Literature Review. **Sustainability**, v. 14, n. 18, 2022. https://doi.org/10.3390/su141811252

NAI, Roberto et al. Public procurement fraud detection and artificial intelligence techniques: a literature review. In: **Companion Proceedings of the 23rd International Conference on Knowledge Engineering and Knowledge Management**. CEUR-WS, p. 1-13, 2022. https://iris.unito.it/bitstream/2318/1888876/1/KM4LAW22_preprint.pdf

OLIVEIRA, G. P. et al. Ranking Public Tenders based on Fraud Alerts. In: **Proceedings of the XII Brazilian Workshop on Social Network Analysis and Mining**. SBC, 2023. p. 1-12. https://doi.org/10.5753/brasnam.2023.232105

REJEB, A.; REJEB, K.; APPOLLONI, A.; KAYIKCI, Y.; IRANMANESH, M. The landscape of public procurement research: A bibliometric analysis and topic modelling based on Scopus. **Journal of Public Procurement**, v. 23, n. 2, 2023. https://doi.org/10.1108/JOPP-06-2022-0031

RIBEIRO, C. G.; INÁCIO, E. The Brazilian government procurement market (2006-2017): Measurement and analysis. Brasília: Institute for Applied Economic Research (IPEA), 2019. Text for Discussion, No. 2476. Available at: <https://hdl.handle.net/10419/211431>. Accessed on: August 2, 2024.

SÁ, T. A.; PESSANHA, J. F. M.; ALVES, F. J. S. Supervised classification methods applied to the identification of supplier fraud. **CONTABILOMETRIA - Brazilian Journal of Quantitative Methods Applied to Accounting, Monte Carmelo**, v. 11, n. 2, p.125-146, 2024.

SCIKIT-LEARN. **RandomForestClassifier**. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html#sklearn.ensemble.RandomForestClassifier. Accessed on: August 2, 2024.

SECOM TCU. **TCU chatbot provides certificates via whatsapp**. 2020. Available at: https://portal.tcu.gov.br/imprensa/noticias/chatbot-do-tcu-fornece-certidoes-pelo-whatsapp.htm. Accessed on: August 2, 2024.

SICILIANI, L.; TACCARDI, V.; BASILE, P.; DI CIANO, M.; LOPS, P. AI-based decision support system for public procurement. **Information Systems**, v. 119, 2023.

SILVA, G. Prevention of irregularities in public contracts: an analysis using Machine Learning techniques. 2022. 91 f., il. Course Conclusion Paper (Bachelor's Degree in Business Administration) - University of Brasília, Brasília, 2022.

SUPREME FEDERAL COURT. **Victor Project advances in research and development to identify themes of general repercussion**. 2021. Available at: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=471331&ori=1>. Accessed on: August 2, 2024.

SUPREME FEDERAL COURT. **STF finalizes tests of new Artificial Intelligence tool**. 2023a. Available at: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=507120&ori=1>. Accessed on: August 2, 2024.

COURT OF AUDITORS OF THE UNION. **Guide to the use of generative artificial intelligence at the Federal Court of Accounts (TCU)**. Disponível em: <https://portal.tcu.gov.br/data/files/42/F7/91/4B/B59019105E366F09E18818A8/Guia%20de%20uso%20de%20IA%20generativa%20no%20TCU.pdf>. Accessed on: August 2, 2024.

VELASCO, R. B.; CARPANESE, I.; INTERIAN, R.; PAULO NETO, O.C.G.; RIBEIRO, C.C. A decision support system for fraud detection in public procurement. **International Transactions Inoperational Research**, v. 28, p. 27-47, 2021. https://doi.org/10.1111/itor.12811

ZAGO, M. F. **State purchasing power as a public policy instrument?** Brasília: National School of Public Administration, 2018.