Death Registry Prediction in Brazilian Male Prisons with a Random Forest Ensemble

Nathan Formentin[1], Eduardo Nunes Borges[1], Giancarlo Lucca[1], Helida Santos[1] , Graçaliz Dimuro[1]

[1]Centro de Ciências Computacionais - Universidade Federal de Rio Grande (FURG)
Av. Itália, km 8 – 96203-900 – Rio Grande – RS – Brazil

{nathanformentin,eduardoborges,giancarlo.lucca,helida}@furg.br

gracaliz@gmail.com

Abstract. *Brazil has the third-largest prison population globally, and it has been growing steadily for more than two decades. Constant growth and low jail investment generated significant problems, such as overcrowding and widespread diseases. This study proposes the construction of a Random Forest classifier to predict the occurrence of deaths in prisons. We extracted data from the National Survey of Penitentiary Information for the years 2015 to 2016. The best-fitted classifier achieved accuracy equals 87% being able to identify correctly up to 84% of deaths occurrences. In the present work, it was possible to establish a relationship between prisons' reality and the data mined, determining areas in need of investment in the penitentiary system.*

1. Introduction

According to the National Penitentiary Information Survey of December 2019, Brazil has 748,009 citizens in prison, of which 37,800 are female inmates, and 710,209 are male inmates [DEPEN 2019], occupying the fifth and third positions in the ranking of largest prison populations in the world, respectively [ICPR 2017]. Besides, one can see the rapid growth of such numbers in the last two decades. In 2000, the number was 232.8 thousand detainees, so there was an increase of 321% in the last nineteen years [DEPEN 2014, DEPEN 2019].

This fast increase in jail population has generated several infrastructure problems in Brazilian prisons, such as overcrowding. The average prison occupation rate in Brazilian states is 161%. The state of Pernambuco has the highest rate, with a prison occupation of 265% of its capacity, and, on the other hand, the state of Maranhão with the lowest prison occupation rate, 121% [DEPEN 2015].

In addition to prison overcrowding, several other factors make Brazilian prisons insalubrious environments, causing the death of citizens deprived of liberty throughout the country. Among those factors, we can mention the war between rival factions [Muggah et al. 2019] and widespread diseases, such as tuberculosis - with an average incidence which is thirty times greater in detainees than in free citizens in the state of Rio de Janeiro, for instance [Sánchez et al. 2008]. Moreover, in Brazilian prisons, there is high vulnerability in the fight against health emergencies, such as the novel coronavirus [Burki 2020] pandemic.

Since 2004, the National Penitentiary Department (DEPEN, acronym in Portuguese) [1] conducts an annual survey involving penitentiaries from all over the country called Infopen [2]. Through a questionnaire, information about infrastructure, human resources, the psychological profile of detainees, and many other factors are obtained. In addition to the visualization of data released by DEPEN, there is also the possibility of obtaining the raw database, which makes it possible to explore the data set and discover valuable information concerning Brazilian prisons, which may be used by competent authorities to make decisions.

Machine learning was already used in prison related context in several occasions, such as text analysis of detainees in different levels of isolation [Becker et al. 2018], prison term prediction based on criminal description [Li et al. 2020] and detainees misconduct problems prediction [Duwe 2020].

Using the data from Infopen 2015 [DEPEN 2015] and the first semester of Infopen 2016 [DEPEN 2016], we propose a classification model applied over the most important features selected to classify if a prison registered deaths or not within that period. The types of deaths analyzed in this study can be accidental, criminal, or due to natural causes, health reasons, or suicide. We used the well-known decision tree ensemble Random Forest [Breiman 2001], since it is considered a state-of-art algorithm in the Explainable Artificial Intelligence field for categorical data. It should be noted that we only considered male and mixed prisons since the number of female prisons was relatively low. Mixed prisons were considered representative since most of their prisoners are male detainees.

We applied feature and hyperparameter selection to optimize the classifier's performance and interpretability. For feature selection, we used the algorithms SelectFromModel (SFM) and Recursive Feature Elimination (RFE) [Chandrashekar and Sahin 2014]. The hyperparameters were optimized by a cross-validated search over a grid (GridSearchCV). These algorithms are available in the library scikit-learn [Pedregosa et al. 2011].

The present work is organized as follows: Section 2 aims to provide information about the data set and the techniques applied for modeling and extracting information from that data. Section 2 is dedicated to the results obtained using different methods for feature selection. Finally, in Section 4, we conclude by analyzing the features considered the most important ones by the different feature selection algorithms applied.

2. Materials and methods

The methodology consists in the application of the process called knowledge discovery in databases (KDD) [Fayyad et al. 1996]. This process aims to extract useful knowledge from a set of data. It is divided into five steps:

1. *selection* consists of choosing the data from which we want to obtain knowledge;
2. *preprocessing* is the application of techniques that eliminate inconsistencies, treat missing data, and provide an initial assessment of the value of specific features for the study;

[1] http://depen.gov.br/DEPEN

[2] http://depen.gov.br/DEPEN/depen/sisdepen/infopen

3. *transformation* consists of preparing data for the application of data mining techniques, such as converting categorical to numeric variables, balancing the data, among other actions;
4. in the *data mining* step, Machine Learning (ML) techniques are applied to the data set to acquire knowledge;
5. *evaluation and result analysis* is the last step. If satisfactory, the process is completed. Otherwise, we go back to any previous phases and adjust the model or the data set to obtain the maximum knowledge in the process.

Figure 1 shows the customization of the KDD methodology used in this work.

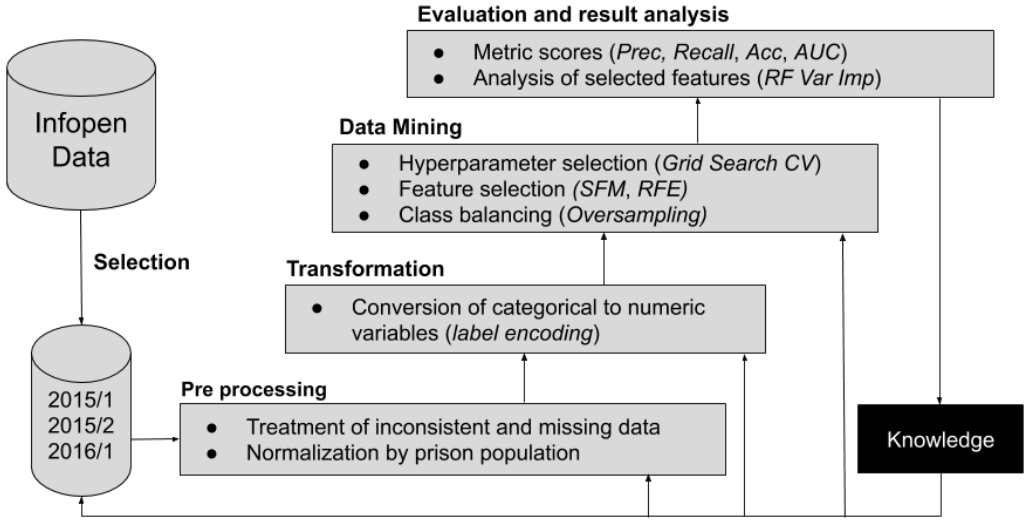

Figure 1. The KDD process applied to discover knowledge in the data of the National Penitentiary Survey.

2.1. Selection

This work used data from the National Survey of Penitentiary Information of 2015 [DEPEN 2015] and the first semester of 2016 [3] [DEPEN 2016]. An online form was used as the collection method, and each prison warden completed it according to the guidelines of DEPEN, the agency responsible for the collection. All data were validated by state managers and also by a procedure performed by DEPEN, named the information consistency check. The data set contains information about infrastructure, management, assistance, inmates' socioeconomic profile, human resources, among other information. A total of 1122 different features were collected. Some of these features are categorical, such as the type of penitentiary. Other features are quantitative, such as the prison population. The total number of instances is 2854, each one representing a prison.

2.2. Preprocessing

The first preprocessing step was to check the possibility of merging the data from the years we selected. To make it easier, we tested whether the applied form followed the same standards, such as the columns' order and the number of features collected. After the join, some data regarded as unimportant were deleted from the data set, such as

[3] http://dados.mj.gov.br/dataset/infopen-national-information-penitentiary

the agent responsible for applying the test (e-mail or name, for example), or the prison location. For storing the data, we used *Dataframes*, a data structure that maintains the data organization as a table, provided by Pandas library [Wes McKinney 2010].

Some columns had many null values, but we found that many of them were multiple-choice questions. Thus, based on the number of "No" and "Not applicable" responses, an auto-complete algorithm was applied to the lines, replacing null values by the most frequent label.Quantitative features that had the number of null values greater than 30% of the number of instances after data treatment (approximately 720) were not considered for the study. The other quantitative features were filled with zero, except for the number of people imprisoned for committing a crime. This decision was made by observing how some of the data were filled in: features about the prison infrastructure not filled in indicate the absence of such resources in the facility. Regarding the variables related to education, considering that only 13% of the detainees in Brazil have access to educational programs [DEPEN 2016], the null values present in these features were also replaced by zero. The quantitative features about committed crimes had their null values replaced by the average value of the feature, since more than 60% of the data comes from public prisons or from institutions that receive detainees who committed several types of crimes.

Incorrect filling of the form was also amended. Wrong data types were corrected. If it was not possible to treat data adequately, the instance was disregarded. Some features had a non-standard filling, generating different labels for the same variable. These characteristics have been fixed or deleted, depending on the number of different responses with the same meaning. Textual categorical features were transformed into numeric variables necessary for the application of ML algorithms. Finally, all instances of the column Status filled with *Incomplete* or *Unavailable* were disregarded.

The main objective of this analysis is to distinguish prisons into two classes, those that registered deaths and those that did not. Therefore, prisons that registered deaths in the analyzed period received the value "one", and those that did not register them received a "zero" value. Prisons that had null values in all variables on deaths were also disregarded.

Concluding the preprocessing stage, all quantitative features, such as the number of prisoners for a particular crime or the number of medical rooms available, had their values divided by the prison population. Moreover, all characteristics that refer to the total prison population were disregarded in the following steps. Those decisions aim to increase the representativeness of the data, for instance, avoiding the pattern that larger prisons, with a higher number of detainees, are more likely to present deaths. Also, it enables one to analyze the availability of specific infrastructure resources for each inmate.

2.3. Transformation

The transformation stage aims to prepare the already preprocessed data set for data mining. Machine learning algorithms deal only with numbers, so categorical variables need to be transformed into numerical ones. In order to do this, we applied the algorithm LabelEncoder [4]. Its objective is straightforward and can be easily exemplified: a column

[4]https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.LabelEncoder.html

with the variables "Yes" and "No" would be transformed into two different numbers: "1" and "0", respectively.

In addition to this transformation, we decided to balance the classes. The number of prisons that did not register any deaths was 1567, and those that registered are 828, or 65.43%, and 34.57%, respectively. Balancing classes is essential to train the classifier correctly since without it, the classifier cannot detect the patterns of the undersampled class accurately. This problem was solved by applying the function RandomOverSampler[5], from the library "imbalanced-learn" [Lemaître et al. 2017]. The function balances the number of classes through a random selection of instances of the less frequent class, with replacement. This new balanced data set was used only for the classifier training, being formed by 3134 instances, 50% of each class.

2.4. Data Mining

After the transformation process, we applied algorithms to select the most important features, aiming to improve the classifier performance and make the results more interpretable. The decision on the number of features was made by extensive testing of the different number of features and hyperparameters. Based on the classifier's evaluation metrics results, we noticed that twenty features were a good number in the trade-off between interpretability and performance. All feature selection algorithms used the importance score of the feature as criteria, which will be explained in section 2.4.2. All the algorithms used for feature selection are available in the library scikit-learn [Pedregosa et al. 2011].

2.4.1. Hyperparameter Selection

Hyperparameters are the model properties that define the learning characteristics. Through them, it is possible to control the model's behavior, generating a significant impact on the performance of the model being trained. For the selection of hyperparameters, we applied an exhaustive search algorithm from the scikit-learn library called GridSearchCV [6]. The algorithm tests all possible combinations of the hyperparameters in the grid, returning the set with the best score in a specific metric. In the case of the present work, we decided to focus on better precision since the other performance metrics were already considered good.

The tested hyperparameters were the number of estimators *(n _estimators)*, which is the number of decision trees that will be generated for decision making, the maximum depth of the trees *(max _depth)*, the criteria for analyzing the node impurity *(criterion)*, and the minimum number of samples for a node to be a leaf *(min_samples_leaf)*.

In addition to these hyperparameters, we applied a cross-validation strategy that consists of dividing our data set for training into n random groups. One division is also defined randomly as a validation group, and the remaining divisions are the training group. This process is repeated until all divisions have been selected as the validation group, and a score is obtained. Stratified cross-validation aims to guarantee the representativeness

[5]https://imbalanced-learn.readthedocs.io/en/stable/generated/imblearn.over_sampling.RandomOverSampler.html

[6]https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

of the data. When specific data are chosen at random, they are not selected again in the next steps. Stratified models reduce bias and variance when compared to non-stratified models [Kohavi 1995]. The algorithm used is available in the scikit-learn library, named StratifiedKFold. Finally, in Table 1, the settings adopted in this study are presented.

Table 1. Hyperparameters tested in the cross-validation grid search and the best values selected using SFM and RFE algorithms.

Hyperparameter	GridSearchCV	SFM	RFE
n_estimators	100,200,300,400,500	300	400
criterion	gini, entropy	entropy	entropy
max_features	auto, *sqrt*, log2	auto	auto
max_depth	5, 10, 15, 20	20	20
min_samples_leaf	10, 30, 50, 70, 90	10	10
n_folds	3,5,10	5	5

2.4.2. Feature Importance

In decision trees, the computation of feature importance consists of how much the node's impurity is reduced, weighted by the probability of reaching the node. This probability is the total of samples where the node is reached divided by the total of samples. A pure node is one where all samples indicate only one answer, while an impure node has no consensus, considering a specific feature. Therefore, we can say that the approximation of the classifier to the purest nodes increases the importance score of the feature, and how representative the feature is to the observations.

Considering that the Random Forest algorithm uses many trees for decision making, we can define the feature importance as the average importance among all trees generated. The features chosen are not necessarily the most important ones, but the group chosen is necessarily the one that most reduces the generated trees impurity. As the data set has quantitative and categorical features, we decided to use the following algorithms, namely: Select From Model and Recursive Feature Elimination [Chandrashekar and Sahin 2014], since they do not need separate methods to deal with the different data types.

2.4.3. Feature Selection

The technique SFM[7] consists of calculating the importance of features and selecting those that score higher than a defined threshold. Another option would be defining the number of features, which we defined as 20. In RFE[8], the feature importance score is calculated for each feature, and then the less important features are pruned. This process is repeated, recursively, until we obtain a data set with a defined number of features, in a step also specified (in our case, in every recursion step, one percent of the data set features are pruned).

[7]https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectFromModel.html
[8]https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html

We used a random forest estimator with the hyperparameters presented in Table 1, obtained through the application of the GridSearchCV algorithm, presented previously. The hyperparameters that are not presented are standard ones for the Random Forest classifier in the scikit-learn library.

The selected features are presented in Table 2. There were selected exclusively by the SFM technique the features: libraries per capita, inmate inclusion from other prisons per capita, toilets for patients per capita, and inmate transfers to other prisons per capita. The features selected exclusively by the RFE model were: brown-skinned people per capita, prisoners homicide per capita, and inmates with disabilities per capita.

Table 2. List of features selected by SFM and RFE algorithms.

Feature	SFM	RFE
Brown-skinned inmates per capita		✓
Dental offices per capita	✓	✓
Existence of a doctor's office in the establishment	✓	✓
Existence of dental office in the establishment	✓	✓
Existence of suture or vaccine rooms	✓	✓
HIV incidence per capita	✓	✓
Incidence of tuberculosis per capita	✓	✓
Inmate inclusions from other prisons per capita	✓	
Inmate transfers to other prisons per capita	✓	
Inmates arrested for extortion involving kidnapping per capita	✓	✓
Inmates arrested for homicide per capita		✓
Inmates with complete higher education per capita	✓	✓
Inmates with disabilities per capita		✓
Inmates with incomplete higher education per capita	✓	✓
Libraries per capita	✓	
Medical consultations held externally per capita	✓	✓
Medical offices per capita	✓	✓
Original inmate inclusions per capita	✓	✓
Pharmacies or medicine stock rooms per capita	✓	✓
Release permits issued per capita	✓	✓
Suture or vaccine rooms per capita	✓	✓
Toilets for health staff per capita	✓	✓
Toilets for patients per capita	✓	

2.5. Evaluation and result analysis

The performance of the fitted models was evaluated using the following metrics [Hossin and Sulaiman 2015]: Accuracy, Precision, Recall, balanced F-measure (F1), and the area under the Receiver Operating Characteristic (ROC) curve (AUC) [Jin Huang and Ling 2005]. For each model fitted, we analysed the relative scores of the top-10 most important features.

3. Results

In this section, we present the data mining findings. After applying the models created in the transformed data set, one can measure how important each feature was for

the decision making of the two models. The features considered most important by SFM and RFE are presented in Figures 2 and 3 respectfully. It is possible to notice that 75% of the features were selected by both techniques. For better visualization, we show only the ten best-ranked features. The graphs present the relative importance, in which 100% is equal to the top-1 feature score of importance: 0.105 for SFM, and 0.1 for RFE model.

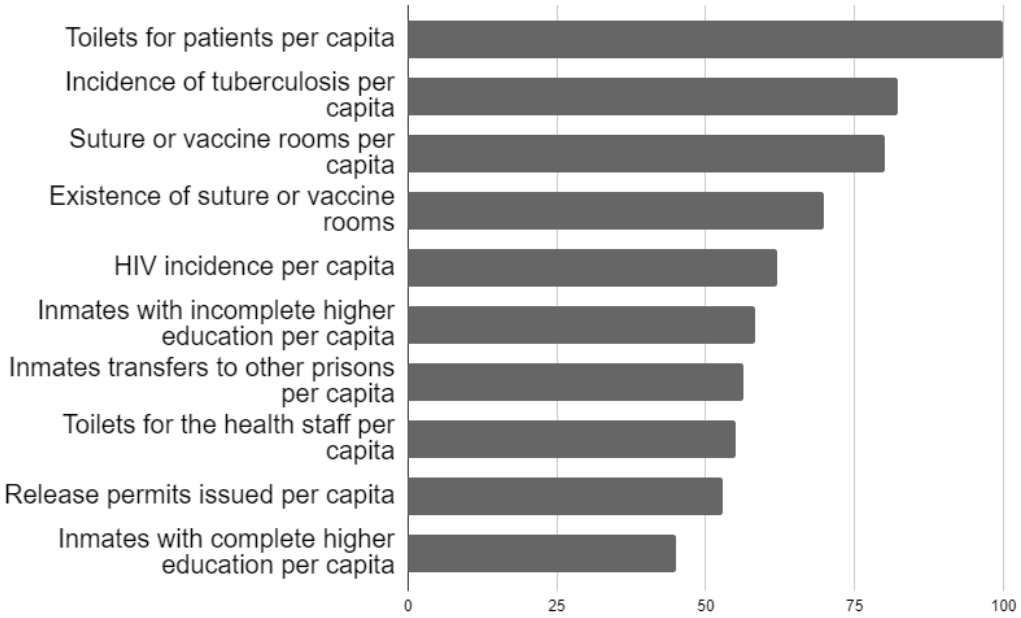

Figure 2. Feature importance score of the selected features using SFM. Values are in percentage, 100% being the value of the best-ranked feature.

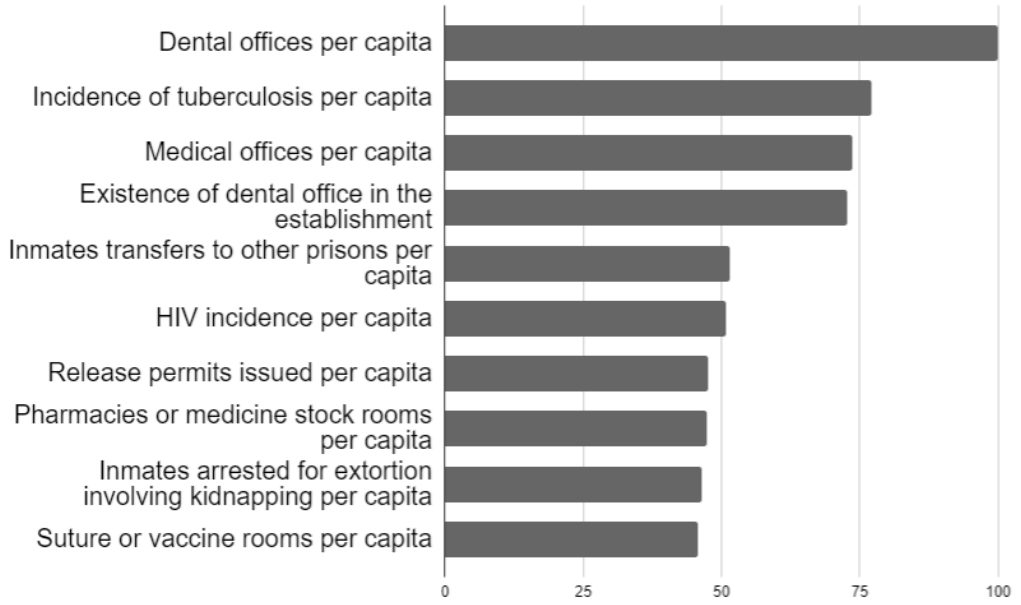

Figure 3. Feature importance score of the selected features using RFE. Values are in percentage, 100% being the value of the best-ranked feature.

Analyzing the four best features selected by the models, it is clear that all of them are directly associated with prisons' health infrastructure. Currently, it is believed

that the health staff provided by the Brazilian government can meet the needs of 200 thousand prisoners, which is less than a quarter of the total number of imprisoned citizens in the country [Soares Filho 2016]. Both classifiers considered relevant for the decision-making the incidence of HIV and Tuberculosis per capita in prisons. Tuberculosis was the most reported disease in Brazilian prisons from 2007 to 2014, while HIV was the third [Miranda et al. 2015]. Both diseases drastically affect Brazil's prisons, being the incidence of tuberculosis in Brazilian imprisoned citizens estimated to be twenty times higher than those who are free [SVS 2014].

In the model where SFM was applied, ranked from the most important to the least important, the other selected variables not presented in the graph were: number of original inmate inclusions per capita, existence of dental offices, inclusion of inmates from other prisons per capita, detainees for extortion crimes through kidnapping per capita, the existence of a doctor's office, medical consultations held externally per capita, libraries per capita, pharmacies or medicine stocks per capita, number of dental offices per capita, and number of medical offices per capita. In the same way, the RFE algorithm selected: number of medical offices per capita, brown-skinned inmates per capita, inmates that committed homicide per capita, number of toilets for medical staff per capita, medical consultations held externally per capita, inmates with incomplete higher education per capita, inmates with disabilities per capita, inmates with complete higher education per capita, existence of a doctor's office, and existence of rooms for suture or vaccine.

For both models, the metric scores of Accuracy, Precision, Recall, F1, and AUC are presented in Table 3. It is possible to see how similar the performances of both fitted models are. But we noticed that the model in which RFE was applied presented a better performance in all classification metrics. The accuracy of the model was 87%, and its area under the ROC curve was 93.5%. Both ROC curves are presented in Figure 4. For prisons with registered deaths, the classifier reached precision and recall equals to 78.9 and 84% respectively. For prisons without deaths, the precision was 91.3%, and the recall 88%. The harmonic mean of the precision and recall (F1) was 89.6%. The presented scores were obtained in the test data set.

Table 3. Classifiers performance metrics scores obtained in the test data set.

	SFM		*RFE*	
Accuracy	0.864		0.87	
AUC	0.934		0.935	
Class	Registered deaths	No deaths	Registered deaths	No deaths
Precision	0.780	0.912	0.789	0.913
Recall	0.830	0.880	0.840	0.880
F1	0.804	0.896	0.814	0.896

4. Conclusion and Future Work

In this study, we proposed two Random Forest models to classify if a prison establishment registered (or not) deaths in a certain period. The high accuracy shows the overall quality of the models. They were able to correctly identify the data class in up to 87% of the instances. The AUC slightly greater than 93% shows the high model's capacity of generalization, which can be adequately applied to unknown data sets.

Figure 4. Comparison between the SFM and RFE model ROC curves.

The selection of the most important features was similar in both models. They presented 20 variables each, with only seven being exclusives. However, the model in which RFE was applied had a slightly better performance in all the metrics evaluated.

After the analysis of the selected features, we concluded that Brazilian prisons present severe problems of infrastructure. Due to overcrowding and low funding, several problems can be noticed, such as the lack of access to health and educational programs within prison establishments.

We could not find any apparent difference between prisons through univariate or bivariate analysis. That is, no direct correlation was detected between the variables so that it was possible to separate the prisons where deaths occurred and those that did not. The lack of apparent differences between the prisons raises the question: how vulnerable are these prisons that have not registered deaths? More research needs to be conducted. Constant improvement in the collection and analysis of these data can be extremely beneficial to decision-makers and competent organizations responsible for the direct funding of the most critical areas.

Future work includes using data from several years provided by DEPEN, which would make viable the study of female prisons. Besides, if more prisons collected and stored information to be studied, it would be possible to thoroughly analyze Brazil's prison problems concerning the present and future issues at national and regional levels. Also, the application of other classifiers would be interesting to evaluate their performance over the data sets used in the present study.

References

Becker, E. J., Burkart, D., Mildner, J., and Tamir, D. (2018). Determination of the defining features of texts written in isolation with a naive bayesian classifier. In *2018 IEEE Integrated STEM Education Conference (ISEC)*, pages 209–210. IEEE.

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

Burki, T. (2020). Prisons are "in no way equipped" to deal with covid-19. *Lancet (London, England)*, 395(10234):1411.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28.

DEPEN (2014). Levantamento nacional de informações penitenciárias – infopen. Technical report, Departamento Nacional Penitenciário, Ministério da Justiça. http://depen.gov.br/DEPEN/depen/sisdepen/infopen/infopen_dez14.pdf.

DEPEN (2015). Levantamento nacional de informações penitenciárias – infopen. Technical report, Departamento Nacional Penitenciário, Ministério da Justiça. http://depen.gov.br/DEPEN/depen/sisdepen/infopen/relatorios-sinteticos/relatorio_2015_2311.pdf.

DEPEN (2016). Levantamento nacional de informações penitenciárias – infopen. Technical report, Departamento Nacional Penitenciário, Ministério da Justiça. http://depen.gov.br/DEPEN/noticias-1/noticias/infopen-levantamento-nacional-de-informacoes-penitenciarias-2016/relatorio_2016_22111.pdf.

DEPEN (2019). Levantamento nacional de informações penitenciárias – infopen. http://depen.gov.br/DEPEN/depen/sisdepen/infopen.

Duwe, G. (2020). The development and validation of a classification system predicting severe and frequent prison misconduct. *The Prison Journal*, 100(2):173–200.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37.

Hossin, M. and Sulaiman, M. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1.

ICPR (2017). World prison brief. Institute for Criminal Policy Research, https://www.prisonstudies.org.

Jin Huang and Ling, C. X. (2005). Using auc and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 17(3):299–310.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, page 1137–1143, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Lemaître, G., Nogueira, F., and Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, 18(17):1–5.

Li, S., Zhang, H., Ye, L., Su, S., Guo, X., Yu, H., and Fang, B. (2020). Prison term prediction on criminal case description with deep learning. *Computers, Materials & Continua*, 62(3):1217–1231.

Miranda, A., Zandonade, E., Job Neto, F., Pompeu, J., Costa-Moura, R., Coelho, R., Saraceni, V., and Fonseca, V. (2015). Análise epidemiológica da situação da saúde

na população privada de liberdade no brasil: dados de bases de informação. *Vitória: Editora da UFES*.

Muggah, R., Taboada, C., and Tinoco, D. (2019). Q&a: Why is prison violence so bad in brazil? *Americas Quarterly*, 2.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.

Sánchez, R., Maria, A. A. M., et al. (2008). *Tuberculose em População Carcerária do Estado do Rio de Janeiro: prevalência e subsídios para formulação de estratégias de controle*. PhD thesis, Fundação Oswaldo Cruz.

Soares Filho, Marden Marques Bueno, P. M. M. G. (2016). Demography, vulnerabilities and right to health to brazilian prison population. *Ciencia & saude coletiva*, 21:1999–2010.

SVS (2014). Situação da tuberculose no brasil. Technical report, Secretaria de Vigilância em Saúde, Ministério da Saúde. http://bvsms.saude.gov.br/bvs/publicacoes/panorama%20tuberculose%20brasil_2014.pdf.

Wes McKinney (2010). Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61.