Exploring bias analysis on judicial data using machine learning techniques

SILVA, Bruno Santos F. and DA COSTA ABREU, Marjory
<http://orcid.org/0000-0001-7461-7570>

Exploring Bias Analysis on Judicial Data using Machine Learning Techniques

Bruno Santos F. Silva
PPgSC/DIMAP
UFRN
bruno@ppgsc.ufrn.br

Marjory Costa-Abreu
Computer Science
Sheffield Hallam Univ.
marjory@shu.edu.uk

***Abstract*—The use of data driven automation is not new, but it has gain a lot of attention recently with the wide-spread understanding that it is the solution to all problems in terms of 'fair' and 'non-bias' classification. This is not different in the law area, where 'artificial intelligence' became a 'magic word'. However, using historic data is a very tricky job which can quite easily propagate discrimination in a very efficient way. Thus, this work is aimed to analyse data from legal proceedings looking for evidence related to the occurrence of bias in the judges' decision-making process, considering mainly the gender or social condition of the convicts. Supervised and unsupervised machine learning techniques, preceded by data analysis and processing procedures, were used to explain and find explicit data behaviour. Our results pointed to the fragility of the techniques to identify biases but suggest the need to improve data pre-processing and the search for more robust classification techniques.**

***Index Terms*—bias, data mining, machine learning, unsupervised learning, supervised learning**

I. Introduction

Although Brazilian courts have acted to digitalise their lawsuits cases, exchanging paper into digital datastores, increasing their workforce and productivity, litigation time are still growing as in [1]. There are currently around 80 million lawsuits in Brazil as reported in [2] and, despite electronic virtualisation of the judicial process covers already a significant part of those cases (around 90%), there is still a considerable challenge to streamline the process, analyse and judge each of the cases individually.

Meanwhile, machine learning has evolved considerably in intelligent data processing techniques [3], [4], being used in so different perspectives of human life nowadays and granting promising applications for generalisation of knowledge. The general understanding is that the application of intelligent techniques as a facilitator of the decision-making process of judges of Brazilian judiciary could increase the productivity of Courts [5] and would give, in theory, more fairness in cases trials.

But, using data based techniques to build decision-making automation could be accompanied by undesired bias, caused by natural changes of legislation or event in socially retired behaviours.

This work describes the analysis and mining on data extracted from judicial systems of the Rio Grande do Norte Court of Justice (TJRN), focusing on looking for data behaviour and correlations, decisions bias made in cases, while checking algorithmic performance and their problem suitability through the application of supervised and unsupervised machine learning techniques [4], [6].

II. Machine Learning and Law

The area of computer science has grown extremely fast in the last fifty years with an exponential speed in the last twenty years because of the advances of the miniaturisation of hardware, the popularisation of the personal computer and more recently with the smartphones/tablets/wearable [7].

But, the 'no understanding' that the general population has regarding technology can create a sense that all that is done in smartphones, computers and automatic solutions is a 'magical' and simple process, which is obviously untrue [8]. And solutions built using Artificial Intelligence techniques are among those confusing and apparently magical cases.

Machine learning is a subset of Artificial Intelligence dedicated to study and develop methods to allow "programs to automatically adjust their performance in accordance with their exposure to information in data" [9]. Those methods work to build mathematical models from analysed and processed data, in a predefined training strategy (which is called learning method), so the models could represent the knowledge in a specific problem domain. This learning is achieved via "a parameter-based model with that are automatically adjusted according to different performance criteria" [9].

There are three types of machine learning methods broadly used: supervised, unsupervised and semi-supervised learning. The supervised concept is related to the presence of a characteristic goal in the dataset (called class), which could be used as a parameter to define what is correct (or not) in machine inferences over the data. Unsupervised learning is based on a dataset where there is no such variable to predict so the challenge is to map relations inside data. And semi-supervised learning techniques are based on labelled and unlabelled data, thus with specific training methods [10].

In a broad review of solutions developed and in use on Brazilian Courts [11], it is notorious the effort being pursued by institutions to improve electronic process-cases systems, in a mix of automation and intelligent solutions mainly focused on dismiss repetitive tasks and improve performance metrics of cases analysis. Most cases use supervised algorithms to perform classification in different problems and perspectives, but there is also a few initiatives using unsupervised machine learning to characterise and discover data relations.

This work have used intelligent classification algorithms to analyse occurrence of explicit bias related to gender and social condition in judicial decisions, followed by applying clustering analysis of dataset to confirm implicit bias related to the same problem classes suggested for classification strategies.

A. *Used techniques*

For the task of classification some supervised machine learning techniques were used in this research, selected because of their ability to generalisation in broad-spectrum problems and increasing complexity of models. There were used k-Nearest Neighbours, Naive Bayes and Decision Trees.

The **k-Nearest Neighbours** is a supervised machine learning algorithm used for classification, is characterised for its independence of data distribution and also of a training cycle for working. It is based on input observations for prediction and a subset of data that have their classes already set, deciding about each data-point based on how similar they are. The main goal of knn is then to define the predict class for a data point based on its distance from k nearest neighbours with most frequent and known classification.

The **Naive Bayes** algorithm works as a statistical classifier and it is based on Bayes theorem of probability considering irrelevant the correlation between features. It is a straightforward, simple, fast classification algorithm, suitable for large datasets as its accuracy and reliability.

Decision Tree is one of the most used supervised machine learning algorithms for classification, mainly because of their ability to clearly represent reasoning strategy, providing good understanding about how decisions are taken and allowing accountability for the model, but also for their flexibility in different problem domains. These are essential issues when managing solutions for Law systems. It works based on a strategy of selecting, in recursive steps, the relevance of the dataset attributes, using Information Gain or Gini Index (methods for identify features relevance), and placing the most relevant one in the root node of the sub-tree, splitting the training set of dataset based on values of chosen attribute.

Clustering (or data mining or unsupervised learning) are techniques which aim to group data that share similarities, and therefore, highlighting data similarities and differences. These models can be built by calculating distances between characteristics and grouping them around estimated centres or hierarchically [12].

In clustering algorithms the goal is to group a set of observations in a sense that ones in the same cluster share more similarities each other than to objects in other clusters.

For this research we have used this techniques: **k-Means** [10] which is based on finding best k clusters where each observation in dataset belongs to one cluster based on minimise distance to their corresponding centres, and **Hierarchical Clustering** which is a method of identifying sets of samples that are similar to each other, but not based on previous definitions (as k from k-Means).

III. Methodology

Machine learning were used in this work because of their computational ability "to synthesise the underlying relationships among data and information" [13]. Data is the main source of inference for our research as we had no strong experts availability to working on it.

Using classification algorithms we should look for behaviour and performance of intelligent models when processing data and prove (or remove) the thesis of existent biases in decision making being affected by gender or social condition of the accused in criminal lawsuits. As a complimentary analysis, but yet central, clustering algorithms processing should be used to clarify data partitioning having each class under investigation as a landmark and reinforcing the bias thesis.

Machine learning techniques suggest some compliance to law-labour model, which involves data classification, patterns identification and decision making [14] based on previous examples and previously classified standards, bringing these technologies into a convincing ally to the management of the judiciary.

A. *Dataset*

The working dataset used for this research was obtained from the "*Além da Pena*" project research group which consists of a scientific action to promote the dialog between the proposed and presupposed law that through empirical investigations studies patterns or interconnections present in the criminal sentences [15], and compiles data from a set of criminal lawsuits from Regional Courts in our State.

It is relevant to reinforce care to the protection of personal data taken when data mining and analysis. Sensitive personal data were not used in this research.

Data attributes include information like year of reference, court ID, gender, sentence penalty of detention or imprison, circumstances considered for establishing sentences, legal aspects used for sentencing, among others, and they were just filled with binary values (as Boolean variables meaning true/false, yes/no or even male/female), a few of them contain string or integer values.

This database was presented to preprocessing methods so machine learning algorithms could handle the dataset. This process included:

1) searching and removing outliers - data wrongly measured or out of reasonable values;
2) removing null rows - data without values in some observations;
3) removing duplicate rows - samples with the same values in all equivalent attributes, inducting particular cases in data;
4) columns discard - attributes with no meaning to study being carried out;

B. *Data Analysis*

In a preliminary step before using machine learning techniques, an exploratory analysis of data was set to allow better understanding about data domain in research, also for analytical and statistical comparison of the sample distribution, observing the original behaviour of the data and looking for preliminary inferences.

For time distribution, we can reach that lawsuits cases are almost homogeneously distributed between 2010 and 2019, but dataset has cases from 2007 until 2020 and this interval has lower number of cases. This could be observed in fig. 1. This means that data used in research is mostly in a regular period basis, with mean value of 52.27 and median value of 54.0 (cases per year), and considering research goals are not time related, this data distribution has minor relevance.

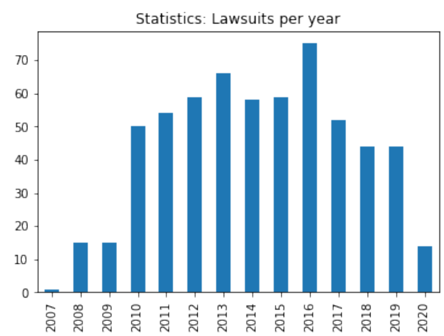

Fig. 1. Lawsuits per year

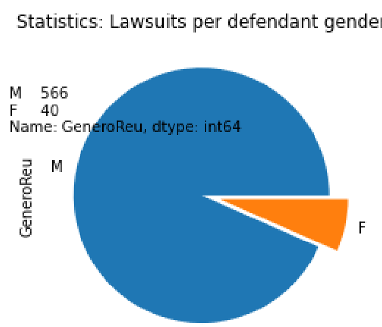

Fig. 2. Lawsuits per Gender

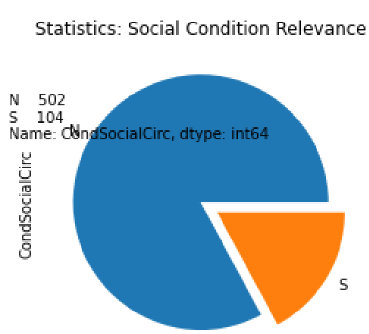

Fig. 3. Lawsuits per Social Condition

The data distribution per gender considering entire dataset presents a imbalanced class attribute, as shown in fig. 2. This attribute is stated in column GeneroReu and illustrates that there are 93.39% of Male accused in collected cases, against 6.6% for accused Female.

Almost the same disproportionality was found in column CondSocialCirc, used for characterise the use of social condition of accused on its judgement. In this case, considering the Boolean values, attribute is in a scale of 17/83, in a measure of 82.83% for cases without use of social condition in penalty delivering. This is shown in Fig. 3.

These characteristics of Beyond the Sentence data is that the chosen classes are clearly imbalanced, such that with these skewed datasets the accuracy metric is not as effective as could be in a problem with a balanced dataset, just because the machine learning algorithms tends to learn more about the most represented class. Therefore, it is important to register that classes have imbalanced data and these cases of source bias should influence negatively in performance of machine learning methods, as we could confirm later in this work.

Using the perspective of heatmap diagram applied to dataset, it is possible to note data distribution related to their values. Heatmap allows analysing dataset from existing values of samples, not with correlation perspective - an usual application of this diagram - and register what may influence the creation of the clusters.

C. Experiments

The analysis of the database Beyond the Sentence, supervised and unsupervised machine learning techniques were used, with complementary and cumulative purposes. Unsupervised techniques, based on unlabelled data, were used for identify characteristics or behaviours in data that were not clearly defined at the beginning of the analysis. On the other hand, the supervised techniques, based on classes explicitly defined prior to data processing, served to confirm whether there would be an underlying classification related to the gender and/or social condition of the accused.

In both cases of machine learning experiments we have used gender attribute of the accused as the first class of analysis (column GeneroReu). For social condition of the accused there is a lack of information in dataset and closest attribute was used in analysis: the column CondSocialCirc could be defined as if social conduct of accused had been judicially considered for penalty delivering.

As part of the initial strategy, data were submitted to unsupervised algorithms without considering the existence of a target attribute, with all attributes being processed regarding the formation of clusters or groups and taking the results for analysis.

In a second approach for supervised algorithms, each class GeneroReu and CondSocialCirc, which represents accused's gender and social condition attributes, were used as model class and the basis for inference of machine learning techniques, leading to confirmation of preliminary goals for our

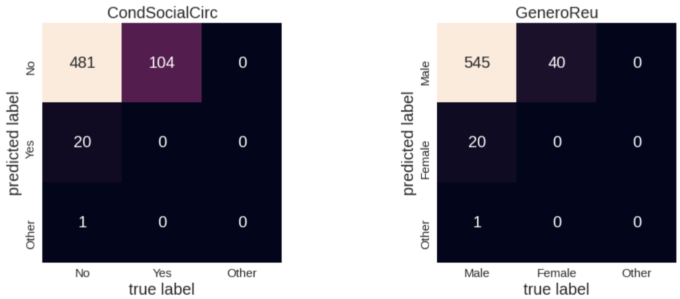

Fig. 4. k-Means results

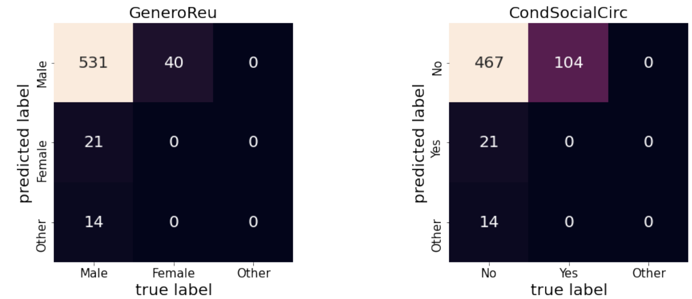

Fig. 5. Hierarchical clustering results

research, occurrence of biases related to decisions taken in lawsuits cases based on gender or social condition of accused.

IV. Results

A. Clustering

When applying k-Means in Beyond the Sentence dataset two complementary strategies were adopted, in the first, all attributes were submitted to processing, without considering that any of them would have meant a classifier for the studied domain. In the second strategy, in different rounds, each class attribute presented above (GeneroReu and CondSocialCirc) was removed from the dataset for training and then used as the target variable of the predictions, allowing the models to be evaluated in relation to their performance and clarify the degree of classification possible for these attributes.

For this step, and using different standard configuration of training algorithm with up to 300 iterations but limited to convergence threshold of 10^{-4}, k-means was submitted with different metrics strategies for comparison in specific dataset, and using performance metrics like execution time, accuracy, Silhouette coefficient, Calinski-Harabasz score, Mutual Information score and Fowlkes-Mallows score, distance metric Chebyshev have major impressive results.

Completing these part of analysis, in fig. IV-A and IV-A are presented the resulting confusion matrices for classes Gender and Social Condition. They reinforce the dominant predictions made in k-Means for Male value and significant amount of false-positives and negatives for the same value, but none for female value. The same behaviour was observed in Social Condition class.

Using agglomerative hierarchical clustering as an alternative comparing method to $k-$Means technique which is based on a bottom-up approach, i.e., the evaluation of clustering formation is made from individual samples up to the groups, arranging them to minimise distance or dissimilarity.

As in $k-$means, the search for a metric suitable for the data domain analysis was carried out by executing the agglomerative hierarchical clustering algorithm over dataset using similar running configurations but changing the metric of distance between the samples and linkage criterion used.

The algorithm executions were repeated 30 times in each configuration, in order to identify variations in the results of the algorithm, being observed average, median and standard deviation from each of the metrics used, being observed no anomalous results (reinforced by minimum values from standard deviation in all executions).

As the dataset is composed of 30 dimensions, the graphic evaluation is impaired for views that broadly illustrate the formed groupings, except for dendrogram diagram that shall be presented further, hence the validation of the quality of the results through appropriate metrics is an essential step in model evaluation.

Agglomerative hierarchical clustering model was build using the same data set applied to $k-$means clustering, but linkage criteria and distance metrics were parameterised in search of the most appropriate configuration for the model and data analysed. Data were submitted to different models based on agglomerative hierarchical clustering to assess the behaviour of the different combinations used, leading to the Chebyshev distance and the criterion of complete and average linkage.

The confusion matrix for the chosen setup is presented in Figure IV-A and reinforce the efficient categorisation of male class with 531 predicted samples as positive-true male, but has no success in female class prediction. Despite that agglomerative hierarchical clustering has predicted 75 samples between false or negative conditions, including a third-class suggested for comparison with k-Means approach with 14 predicted observations.

Similar results were observed in Social Condition class, as presented in Figure IV-A, were predicted values for 'No' use of social condition in sentences had been highly assertive (over 77%), but 'Yes' to use of such condition had no positive predicted value.

In attention to our prime goal in these work, we came to a partial observation that the clustering algorithms based on machine learning do not demonstrate the occurrence of biases for the classes of gender and social condition of accuseds.

Considering that were not used previously defined classes attributes in these algorithms, the occurrence of non-obvious bias have not been demonstrated in these experiments duo to low separability accuracy achieved by the clustering techniques. But some partial occurrence of separability suggests that different models should be analysed for give more accurate results.

B. Supervised learning

Following algorithms were responsible to make binary classification of the accused's gender in the lawsuit data, which can be identified as 'Male' or 'Female', and also binary classification of the use of social condition criteria in sentencing decisions (which have values 'Yes' or 'No').

Data set for training and testing was splitted on a 70-30 proportion for training and testing experiments. These procedure was repeated for both classes, isolating data of column attribute designed as class and using all other columns as subset for further actions over learning procedures.

As k-Nearest Neighbours (knn) is a non-parametric algorithm, i.e., it does no assumption for underlying data distribution, and also is a lazy learning technique - it does need no training data points for model definition, this supervised machine learning was chosen for the initial tests in this research work.

When applying knn algorithm to data set Beyond the Sentence, first challenge is to define a value of k that is low, preferable odd (our classes have binary values) and fitted to data domain. The method used here was to submit samples to training and testing in knn but varying the values of k and computing the accuracy as a comparative measure.

Following premises stated above and according to accuracy variation, where by varying the k it was achieved the best accuracy both for training and testing, suggesting the values of k would be set to 3 or 5. But, after comparing common distance metrics, results shows that $k = 5$ gave better results but similar when used with Euclidean, Manhattan, Jaccard and Minkowski distance metrics.

It is important to note that after each new random split of dataset for training and validation of the model, results vary in a manner that suggest knn has a great sensitivity to data distribution and impact its prediction results.

After defining knn model with $k = 5$ and Euclidean distance metric, the machine learning was tested against classifying GeneroReu and CondSocialCirc classes and the results were compiled in Table I. These data shows that even accuracy is high, classification of less representative categories is minimal (as female gender or no use of social condition in penalty).

TABLE I
PERFORMANCE RESULTS FOR K-NN

Class	ExecTime	Accuracy	F1 Score	MCC
Gender	0.0110	0.95055	0.18187	0.30826
Social Condition	0.0115	0.81318	0.19047	0.24919

These results could be reinforced by confusion matrices presented in Figures IV-B and IV-B which reveals high accuracy classifying Male gender and also for 'no use of social condition in penalty' duo the their filling of data set, in opposite to the very low accuracy related to other classes of each attribute.

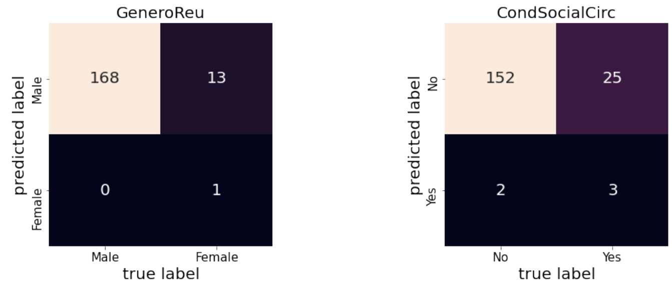

Fig. 6. kNN results

Despite the high accuracy, the results are not so satisfactory, due to the low mcc value, the main metric we are considering. However, its value greater than 0 indicates that the classification of defendants gender in the lawsuit, as Male or Female, or even about the use of social condition in penalties, can still be done and this is a relevant contribution of this work.

Second classification method applied to analyse results in gender and social condition classes was the Naive Bayes supervised algorithm. Naive Bayes algorithm is a statistical technique and has some variants according to identified data set probability distribution. Based on preliminary assumption about classes binary values, the Gaussian Naive Bayes was the chosen variation to experiment. After multiple (30x as default) execution of the algorithm and with variance measures in order of 10^{-17}, results were those presented in Table II.

TABLE II
PERFORMANCE RESULTS FOR NAIVE BAYES

Class	ExecTime	Accuracy	F1 Score	MCC
Gender	0.00089	0.20329	0.15204	0.10809
Social condition	0.00071	0.26923	0.35748	0.08998

Performance metrics for Gaussian Naive Bayes execution shows that execution time is the lowest since then, and accuracy and Mathews Correlation coefficient are poorly significant despite better results on F1-score, which express better performance for recall and precision, specially when submitted for CondSocialCirc class, which is related to social condition of accused.

Additionally, the resulting confusion matrices of execution for the classes are presented in Figures IV-B and IV-B providing evidence on the results achieved. The Naive Bayes predicted almost all observations as being female (in class GeneroReu, and therefore with a high hit rate in this class, and the same behaviour occurred in the social condition class, where the use of this class was being mostly predicted correctly (True-Positive with 97.36%).

This algorithm classified almost all data as 'Female' and as positive use of Social Condition and because of that, its results are much worse than that of K-Nearest Neighbour, as presented above.

The Naive Bayes is based on applying Bayes theorem with strong independence assumptions between the features, also

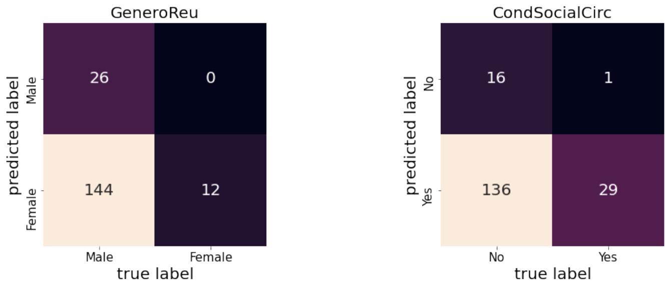

Fig. 7. NaiveBayes results

considering dataset has imbalanced classes in both analysed cases (gender and social condition), this context induces higher probability to majority values and this would influence overall performance of Naive Bayes when submitted to classification task.

The latest classification technique was with Decision Tree algorithm and its definition was based on the comparison between the performance metrics obtained from the standard configuration of library followed by tuning the selection criteria of the best attributes (Gini index or Entropy), the strategy of dividing the data at each iteration and, finally, the depth limitation of the tree (which may influence overfitting or underfitting).

Based on the results achieved after tuning, the model was defined and built using the Gini Index selection criteria, with a strategy for selecting the Best measured attribute and calculating the tree to the maximum required depth.

The decision tree model was then submitted to the data set for 30 different cycles, considering separate experiments for each class under analysis, which resulted in the performance measurements listed in the Table III. Columns are organised with parameters execution time, accuracy, F1 score and Mathews-Correlation coefficient.

The scores achieved confirm the excellent performance of the model based on Decision Tree, with high accuracy in both cases but with emphasis on the prediction of Gender, where it reaches 89.79% of assertiveness. On the other hand, given the weakness of considering only the positive hits, it is in the class of Social Condition the model presents a better general assertiveness, given the F1 and mcc indices with 0.55 and 0.45 respectively, illustrating a reasonable assertiveness in the other quadrants of the confusion matrix.

TABLE III
PERFORMANCE RESULTS FOR DECISION TREES

	execTime	**Acc**	**F1 score**	**MCC**
generoReu	0.001842	0.89798	0.33005	0.27632
CondSocialCirc	0.00175	0.82106	0.55782	0.45585

While the last two tested algorithms showed not so good results, the decision tree was the algorithm that made the best classification, with about half of Female gender classified correctly and almost all Male. Despite having a lower accuracy than the k-Nearest Neighbours, its mcc value was higher, indicating its higher quality in the classification of data in this scenario.

One reason for this is that the algorithm is favoured by the large number of categorical features in the database and it is not as influenced by the dependence between the features, as Naive Bayes model.

V. Final remarks

As this research work was aimed to analyse and mining judicial data looking for identify bias in data related to gender or social condition classes based on dataset available from X project research and, for that, resources were used to build supervised and unsupervised machine learning models, resulting in data analysis, inferences and contributions.

The results achieved with experiments, after procedures for tuning, shows inability of clustering algorithms to isolate suggested classes, related to gender or social condition of accused in criminal lawsuits. This means that related to dataset analysed, non-obvious bias was not confirmed and could not create a separate space between subset of data from different values for proposed classes.

Related to clustering, results achieved also demonstrate that imbalanced dataset has a definitive influence in how algorithms perform groups when trying to isolate observations, their results were biased but related to classes with major presence in dataset and do not imply in good prediction or clustering results. But these results were improved as model's robustness increases, suggesting that classification should be pruned for different techniques.

When analysing supervised machine learning algorithms results and compiled performances reinforces poor efficiency of classification for this proposed context when using gender and social condition as classes for Beyond the Sentence dataset. Considering improved results in Decision Trees in general, there is an indicative to using more complex domain machine learning techniques to review and confirm explicit biases in decisions made.

This results allows us to conclude that, despite no strong evidences of existence of (obvious) bias in judge's decision making in sentencing related to gender or social condition in analysed dataset and collected features, there is field to keep on research and look for the occurrence of bias in such problem domain.

For improved results, would be necessary to use different approaches in data preprocessing and when dealing with small imbalanced dataset, to inference, like using written sentences and more informational features within this space problem.

Comparing results achieved using unsupervised and supervised machine learning techniques have demonstrated there is evidence of better performance when using classification algorithms than clustering. Considering the occurrence of imbalanced classes in analysed dataset suggest that training phase with knowledge of data behaviour, step not available in unsupervised techniques, allows better performance results to classifiers.

Another important evidence from this research work is related to the relevance of use of multiple algorithms, distances and performance metrics, tuning and evaluating the results while looking for the best model for the problem. Defining a model in machine learning depends on this try-analyse approach.

References

[1] Conselho Nacional de Justica. ”Inteligência Artificial no Poder Judiciário Brasileiro”, Coordenação: José Antônio Dias Toffoli, Bráulio Gabriel Gusmão, Brasilia, 2019.

[2] Conselho Nacional de Justiça, ”Justiça em Números - Analytical Report 2020”, baseline 2019, Brasilia, 2020.

[3] S. Haykin. ”Redes Neurais: princípios e prática”. 2nd ed, Bookman Editora, 2001.

[4] Goodfellow, I. and Bengio, Y. and Courville, A. ”Deep Learning”. MIT press, 2016.

[5] Coelho, Joao Victor A.B.R., ”Aplicacoes e implicacoes da inteligencia artificial no Direito”. BCZM, UFRN, 2017.

[6] Branting, K.L. ”Automating Judicial Document Analysis”. In Proceedings of the Second Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL 2017), CEUR, London, UK, 2017.

[7] Boonstra, T.W. and Werner-Seidler, A. and ODea, B. and Larsen, M.E. and Christensen, H. ”Smartphone app to investigate the relationship between social connectivity and mental health”. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 287–290, 2017.

[8] Augustine, N. ”What we don’t know does hurt us. How scientific illiteracy hobbles society”. In Science, vol 279, number 5357, pages 1640–1641, March 1998.

[9] Laura Igual and Santi Segui. ”Introduction to Data Science - A Python Approach to Concepts, Techniques and Applications”, Springer, 2017.

[10] Grus, J. ”Data Science from Scratch”. O’Reilly Media, 2015.

[11] Da-Costa-Abreu, M. and Silva, B.S.F. ”A Critical Analysis of ’Law 4.0’: The use of Automation and Artificial Intelligence and their Impact on the Judicial Landscape of Brazil”. In Revista de Direitos Fundamentais e Tributacao, Grupo de Pesquisas Avancadas em Direito Tributario - PPgD/PUCRS, 2020.

[12] Linden, R. ”Técnicas de Agrupamento”, Revista de Sistemas de Informação da FSMA, number 4, pages 18–36, 2009.

[13] Mariette Awad and Rahul Khanna. ”Efficient Learning Machines - Theories, Concepts and Applications for Engineers and System Designers”. Apress Open Media, 2015.

[14] Aikenhead, M. ”Uses and Abuses of Neural Networks in Law”. Santa Clara Computer & High Tech. LJ, vol 12, page 31–55, HeinOnline, 1996.

[15] Saboya, Keity. ”Ne Bis In Idem em Tempos de Multiplicidades de Sanções e de Agências de Controle Punitivo”. Jornal de Ciências Criminais, vol 1, pages 71–92, 2014.