Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 30, Num. 02 (2023) 51-58

RESEARCH ARTICLE

Catboost algorithm application in legal texts and UN 2030 Agenda

Aplicação do algoritmo Catboost em textos jurídicos e na Agenda 2030 da ONU

Lucas José Gonçalves Freitas[1,2]*, Pamella Sada Dias Edokawa[2], Thaís Carvalho Valadares Rodrigues[1], Ariane Hayana Thomé de Farias[3], Euler Rodrigues de Alencar[2]

Abstract: This article evaluates the application of the Catboost algorithm for automatic classification of legal texts in The United Nations (UN) 2030 Agenda for Sustainable Development Goals (SDGs). The task consists of labeling texts from initial petitions and rulings based on identifying topics related to the objectives of the 2030 Agenda, which include sustainable development, quality education, gender equality, preservation of the environment, among other topics of interest to UN member countries. This work aims to help Judicial System employees in case management task, an activity that is manual and repetitive. Since the Catboost algorithm allows joining textual, numerical and categorical features in the same classification model. The proposed approach adds to the classification algorithm traditional metadata about legal processes, such as the Supreme Court Class and Field of Law. The main contributions of this work are: analysis of metadata in machine learning flows and evaluation of the Catboost algorithm for textual classification in legal contexts.
Keywords: Natural Language Processing — Legal Text Classification — Machine Learning — UN 2030 Agenda

Resumo: Este artigo avalia a aplicação do algoritmo Catboost para classificação automática de textos jurídicos em objetivos de desenvolvimento sustentável (ODS) da Agenda 2030 da Organização das Nações Unidas (ONU). A tarefa consiste em utilizar textos de petições iniciais e acórdãos com base na identificação de assuntos relacionados aos objetivos da Agenda 2030, que incluem desevolvimento sustentável, educação de qualidade, igualdade de gênero, preservação do meio ambiente entre outros temas de interesse para os países membros da ONU. O objetivo deste trabalho é auxiliar servidores do Poder Judiciário na separação ou agrupamento gerencial de processos, atividade que é manual e repetitiva nos órgãos públicos. O algoritmo Catboost permite reunir variáveis textuais, numéricas e categóricas no mesmo modelo de classificação e a aplicação proposta acrescenta, no processo de classificação, metadados tradicionais sobre processos jurídicos, como a classe processual e o ramo do direito. As principais contribuições deste trabalho são: análise dos metadados em fluxos de aprendizagem de máquina e avaliação do algoritmo Catboost na atividade de classificação textual em contextos jurídicos.
Palavras-Chave: Processamento de Linguagem Natural — Classificação de Textos Jurídicos — Aprendizagem de Máquina — Agenda 2030 da ONU

[1] *Departamento de Estatística, Universidade de Brasília (UnB), Brasília - Distrito Federal, Brasil*
[2] *Secretaria de Gestão Estratégica, Supremo Tribunal Federal (STF), Brasília - Distrito Federal, Brasil*
[3] *Corregedoria Geral de Justiça, Tribunal de Justiça de Roraima (TJRR), Roraima - Boa Vista, Brasil*
*__Corresponding author__: lucas.freitas@stf.jus.br
DOI: http://dx.doi.org/10.22456/2175-2745.128836 • **Received:** 12/12/2022 • **Accepted:** 12/04/2023

1. Introduction

The Federal Supreme Court (in Portuguese, Supremo Tribunal Federal - STF), the highest instance of the Brazilian judicial system, produces an immense amount of data, usually organized in text form, through decisions, petitions, injunctions, appeals and other documents. Text classification is a fundamental part of the assessment stage of lawsuits cases. This step, when done well, serves to organize information from legal actions and to retrieve historical information from judgments and decisions, for example. In this context, a supervised learning tool for classifying legal documents into 17 goals (SDGs) of the UN 2030 Agenda for Sustainable Development can be of great use to the court, since this is manually performed by several specialists nowadays.

The objective of this work is to use the Catboost algorithm, which gathers text features and metadata, to automatically label legal pieces in Sustainable Development Goals of the UN 2030 Agenda. In this work, the metadata chosen were: keywords count for each SDG, Supreme Court class (Classe Processual, in Portuguese) and field of law (Ramo do Direito, in Portuguese). The SDG keywords are collections of words related to each SDG of the 2030 Agenda, which are produced by specialists in the legal field of the court. Supreme Court class and field of law are typical metadata available upon process registration and are strong indications of the nature of the judicial process. Supreme Court Class (STF class) and STF actions are synonymous in the present work.

Recently, some international works provided great impact in the area of legal text classification. In Katz, Bommarito and Blackman [1] they used Random Forest algorithms to predict US Supreme Court decisions. Medvedeva, Vols and Wieling [2] present a method based on support vector machine (SVM) models to predict decisions of future cases in the European Court of Human Rights. Hausladen, Schubert and Elliot [3] presented machine learning applications to predict liberals and conservatives decisions in several North American courts. Radygin et al. [4], developed an artificial intelligence tool to search for violations in Russian Federal Legislation. In Brazil, also in the law and court context, there are the works of Junior, Calixto and Castro [5], Nascimento and Oliveira [6], Menezes and Clementino [7] and Zanuz and Rigo [8], where the authors work, respectively, with ontologies, clustering, transformers architecture and recognition of named entities.

The main contributions of this article are an analysis of metadata in machine learning flows (SDG keywords, STF Class and Field of Law) and an evaluation of the Catboost algorithm in the activity of textual classification in legal contexts, given that many other algorithms have already been tested and do not allow to combine text and numerical and categorical features.

This article is organized as follows: Section 2 revisits the UN 2030 Agenda, the Section 3 deals with the database used. Section 4 presents theoretical aspects of the Catboost algorithm (methods) and the Section 5 brings results and final considerations.

2. UN 2030 Agenda

The UN 2030 Agenda is an action plan for people, the planet, prosperity, universal peace and freedom. Created in 2015, this action plan, which involves objectives such as eradicating poverty, reducing inequality, preserving the environment and sustainable economy, is being organized into 17 SDGs (Sustainable Development Goals), as can be seen in [9]. The 17 SDGs are integrated – action in one area will affect results in others – but the texts of each SDG are completely different in essence, so identifying them them as labels in legal texts is an important step towards automating repetitive tasks and improving workflows. Including the UN 2030 Agenda in the court routine is a strategy to make justice more efficient, as it is not important just to judge quickly, judicial decisions must positively affect people's lives. In practice, labeling legal actions or texts according to the UN 2030 Agenda can be a way of anticipating judgments and decisions with important social impact, better serving the population as a whole.

The Federal Supreme Court has a hotsite [10] where information about the Agenda 2030 is available, as well as a panel containing the metadata of the processes currently labeled with SDG in court. These data can be easily downloaded in .xlsx (Excel) format there. Figure 1 shows examples of the Agenda 2030 SDGs.

Figure 1. Examples - Agenda 2030 SDGs.

Experts can identify UN 2030 Agenda ideas in parts of legal texts in many ways. A simple and easy-to-use method, however, does not depend on extensive legal experience. This is the active search for representative words of each SDG, here called SDG keywords. When observing, for example, the term "*endangered species*" in a given legal text, one immediately sees its relationship with the nature and environment SDGs, such as SDG 14 (Life below water) and SDG 15 (Life on land). Other examples of SDG keywords will be given in Section 3, which is dedicated to the database. Such identification greatly facilitates the separation of cases by theme, for example. This justifies the inclusion of keyword SDGs created by specialists in the flow of aid for text classification via machine learning. Since this is a search made by regex and automation, there is an immense versatility. It can be used in texts of any nature and be easily implemented in several programming languages, as long as experts in the field in question indicate the most important technical terms for the search.

3. Database and Text Processing

3.1 Data Base

The database contains 1643 petitions and rulings duly labeled with the UN 2030 Agenda SDGs. This number of processes is substantially smaller than the amount of table entries available on the hotsite, as not all tagged processes have initial requests or decisions available and legible. Repeated processes and documents whose text is not readable, as well as records with less than 15 words, were excluded from analysis. The adjective readable applies to processes with parts in native digital

PDF or properly scanned by Optical Character Recognition (OCR Tesseract), a task that extracts text from images through Long Short Term Memory (LSTM) neural networks.

The input table for machine learning contains the process ID, a binary column for each SDG indicating label (1) or not (0), keyword count for each of the SDG (keywords are highly representative terms of each SDG), the Supreme Court class and the field of law, in addition to a column dedicated to the clean text, whose cleaning will be better explained in the subsequent sections. STF actions refers to the procedure adopted in the judicial sphere, such as extraordinary appeals, claim of noncompliance with a fundamental precept and many others. Field of law refers to the main areas of legal practice, such as civil law, criminal and environmental law.

In Brazil, there is a general data protection law, called LGPD (in Portuguese, Lei Geral de Proteção de Dados Pessoais). This law indicates that personal data must be treated with due confidentiality by companies and public bodies. The data used in the present work are public and, in full compliance with the LGPD, are available on the STF hotsite for Agenda 2030 and on the court's own page, in the procedural search link. The initial manipulation of the database available on the court's hotsite and the reading of the PDF texts are done in R language, while the Catboost family models were adjusted in Python language, hence the need to generate a parquet file with the base. The texts have different sizes, but they are all considered long in relation to texts extracted from social networks or news sites. Figure 2 shows the distribution of STF actions tag.

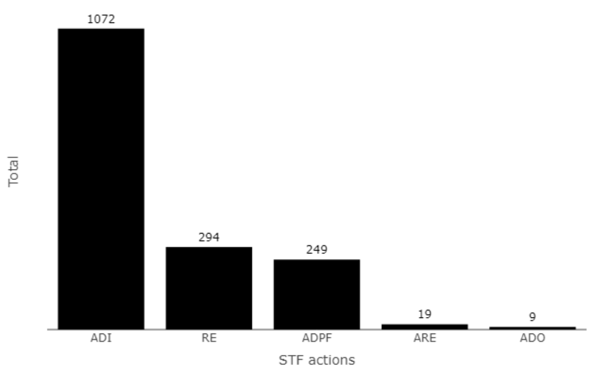

Figure 2. STF class of labeled actions.

It is possible to see that the ADI class corresponds to more than 50% of the database. Whereas Figure 3 highlights that he field of law most frequent in the Administrative Law. There is a clear imbalance between the STF actions and the fields of law of the cases in the sample, but this situation represents well the process receipt flow the court, which is spontaneous and does not keep the same proportions for the process metadata.

In the final considerations, specific results will be presented for action with ADI class and field of Administrative Law and other Public Law Matters. The objective is to evaluate whether there is an increase in the accuracy of Catboost

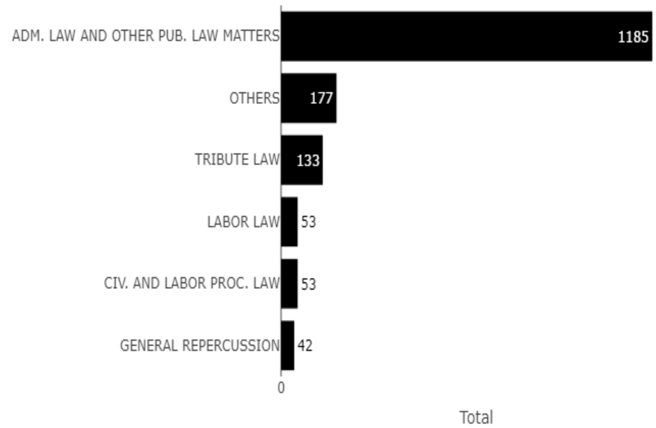

Figure 3. Field of Law of labeled action.

models using only processes with this class and field.

Keyword count is a feature built on the experience of numerous servers at the court classification sectors. The keywords are strong indicatives of the 2030 Agenda SDGs. The word "*feminicide*", for example, indicates that a certain legal piece is likely to be related to SDG 5 - Gender Equality. Another example is given by the occurrence, in large frequency, of the word "*vaccine*", which normally relates the text to SDG 3 - Health and Welfare. It is not reasonable to use only keywords to label a process, given that in Portuguese words have different meanings, depending on the context. Moreover, a word can appear a lot in a specific part of an initial petition, for example, just because its author wants to use another piece of example or basis for his request. Word count will be used in the Catboost as a model numerical feature, instead of being a rigid decision rule for the classification. In the present work, the word count features were constructed in a binary way, with the number 1 indicating the presence of one or more keywords and the number 0 indicating the opposite. This construction can be reviewed in future works.

3.2 Text processing

Text processing step is a very important step for Natural Language Processing (NLP), as it can reduce the complexity of the problems and also impact models performance. In this work, we used the text cleaning flow presented next:

- Removal of stopwords and punctuation from the Portuguese language;
- Removal of special characters such as #, @ and &;
- Removal of unnecessary numbers;
- Removal of unnecessary legal terms;
- Removal of unnecessary whitespace;
- Lowercase of all text words and post tag (optional).

Table 1 shows an example of raw text and its corresponding text cleaned by the processing step.

Many text-cleaning approaches in natural language processing also use tokenizations and lemmatizations, which are ways of reducing words to radicals, with or without context. These strategies serve to reduce the number of words in the text dictionary (bag of words - BoW) and, consequently, computational cost. In this work, however, the option was to compare basic cleaning (the list above) and an even more aggressive cleaning, using only words classified as nouns, adjectives, adverbs and verbs. The idea is to verify which type of cleaning performs best in legal texts, especially in the context of case classification.

Post tag activity, from the English part of speech tagging, is a kind of grammatical classification of words in a text. It is a step that can be intermediate, as in the case of this work, or the task of interest in NLP flows. The most common tags are ADJ (adjective), ADV (adverb), NOUN (noun), PROPN (proper nouns) and VERB (verbs), which are usually standardized to allow their use in multiple languages. The post tag step can be done, in Portuguese language texts, through the Spacy package [11] in Python (see Walsh [12]). The Spacy framework has excellent pipelines in Portuguese, trained for tasks such as post tag, named entity recognition, tokenization and lemmatization. With appropriate adaptations, it can also be used to recognize text keywords, that is, to extract the most important words from each text.

4. Methods

4.1 Catboost Algorithm

The Catboost models are machine learning algorithms that use gradient boosting and decision trees, they were proposed by Yandex in 2017 [13]. The classifiers ensemble, which are traditionally divided into boosting and bagging, are sets of classifiers trained individually and combined in some way, the most popular being given by votes [14]. In the boosting approach, several classifiers (decision trees) are trained sequentially and adaptively, where the current state model depends on the previous models and tries to improve them. Differently from the bagging approach, where the models are trained in parallel and the voting is done in a simple way, the votes in the boosting approach are weighted by the performance of each model. Figure 4 shows the difference between these two ensemble approaches.

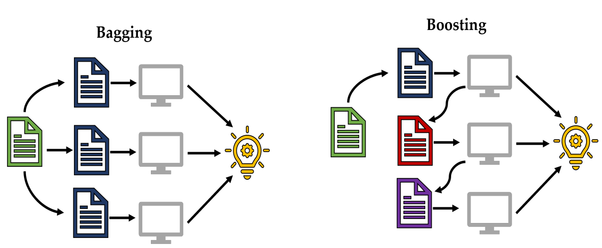

Figure 4. Bagging vs Boosting strategies.

One of the main differences between Catboost and other boosting algorithms is the native support for categorical data, which is transformed into numerical data through formulas available in the documentation [13] or via one-hot-encoding. The XGBoost algorithm, for example, only accepts natively numeric variables. Table 2 presents a toy example with categorical variables. Considering a possible classification task centered on the label variable, the first step of the transformations performed by the Catboost algorithm is to combine the categorical features, in this example given by the genre and subgenre columns. The new base is given in Table 3. Next, genre column will be transformed into a numerical variable using the following formula, which is one of several options available in the Catboost algorithm documentation:

$$AVG_i = \frac{count + prior}{total + 1}, \tag{1}$$

where:

- $i = 1, 2, 3 \ldots N$, with N = sample size;
- $count$ = number of occurrences of label 1 in previous entries of the same category;
- $prior$ = parameter defined a priori. Here, prior = 0.05;
- $total$ = number of previous objects of the same category.

Applying the proposed transformation and exchanging genre column for AVG vector, we have the results presented in Table 4. Therefore, AVG column is just a numerical and alternative way of representing categorical genre features. In the Catboost algorithm, the numerical features can be normalized or not, depending on the analyst's choice. Text features are handled by BoW (Bag of Words) and other dictionary approaches. As an example of the text handling done by Catboost, Table 5 shows a text feature. Here, the texts are broken into the smallest possible unit, the words, as show in Table 6. A dictionary of words is created, along with an identification column for each word in the dictionary (Table 7). Finally, texts are transformed into numerical resources, as shown in the Table 8.

In addition to the differences in the treatment of categorical and text features, there are significant changes in the tree classifiers that make up the ensemble. The Catboost algorithm uses symmetric trees (oblivious trees), which are balanced trees and less prone to overfitting, as they use the same splitting criterion at all tree levels. Other boosting algorithms use asymmetric trees, as in LightGBM (Light Gradient Boosting Machine). Another significant difference occurs in the use of gradient boosting. Some boosting models, such as XGBoost and LightGBM, use all points in a sample to train a model and calculate its residuals. Subsequently, the same sample is used, with the target residuals, in another model. The intention is to make the residuals smaller and well behaved. This process is repeated for several iterations and is clearly susceptible to

Table 1. Raw and Clean Texts.

Original text: supremo tribunal federal ementa e acórdão inteiro teor do acórdão - página 1 de 6 09/12/2014 primeira turma ag.reg. em tutela antecipada na ação cível originária 1.824 amapá relator : min. marco aurélio agte.(s) : estado do amapá proc.(a/s)(es) : procurador-geral do estado do amapá agdo.(a/s) : união proc.(a/s)(es) : advogado-geral da união pessoa jurídica de direito público
Clean text: primeira turma tutela antecipada cível originária amapá relator marco aurélio estado amapá procurador geral estado amapá união advogado geral união pessoa jurídica direito público

Table 2. Example - Categorical feature.

order	genre	subgenre	label
1	metal	black	0
2	classic	opera	1
3	metal	trash	1
4	metal	black	1
5	metal	trash	0
6	classic	concert	1
7	classic	concert	1
8	metal	trash	0

Table 3. Categorical feature (Continuation).

order	genre	label
1	metal black	0
2	classic opera	1
3	metal trash	1
4	metal black	1
5	metal trash	0
6	classic concert	1
7	classic concert	1
8	metal trash	0

Table 4. Example - Categorical feature (Final.).

order	AVG	label
1	0.050	0
2	0.050	1
3	0.050	1
4	0.025	1
5	0.525	0
6	0.050	1
7	0.525	1
8	0.350	0

Table 5. Example - Text feature.

order	text
1	Cats are so cute!
2	Rats Scare me!
3	The cat catches the mouse
4	How cute! Mice and Cats Playing
5	An army of mice runs after the cat
6	The cat asks the mouse for peace
7	The cat is afraid
8	Cat and mouse live in peace

overfitting, given that the same data are used in the initial training and in the improvement of the residuals. The Catboost algorithm, in turn, uses a technique called ordered boosting. In practice, it is a method of obtaining residuals only with data not yet used, and for this, it is necessary to impose an arbitrary order on the input data. Next, evaluated scenarios and models performances will be presented. Algorithm implementations are available at [15].

4.2 Evaluated scenarios

The present work intends to compare the performance of the Catboost algorithm in different scenarios, namely: with or without post-tag and with all possible combinations of texts and categorical features. The following list shows all scenarios:

- Scenario 1: STF Actions;
- Scenario 2: Field of Law;
- Scenario 3: SDGs keyword Count;
- Scenario 4: STF Actions and Field of Law;
- Scenario 5: STF Actions and SDGs keyword;
- Scenario 6: Field of Law and SDGs keyword;
- Scenario 7: STF Actions, Field of Law and SDGs keyword.

For each scenario, full text and text with pos tag options will be considered. The main objective of this work is to verify if the Catboost algorithm can reach a good accuracy in the prediction of SDG tags, additionally pos tag performance will also be discussed, as legal texts can be large, which increases the complexity of NLP algorithms. Having one more viable option to decrease records with quality can be interesting for data processing flows.

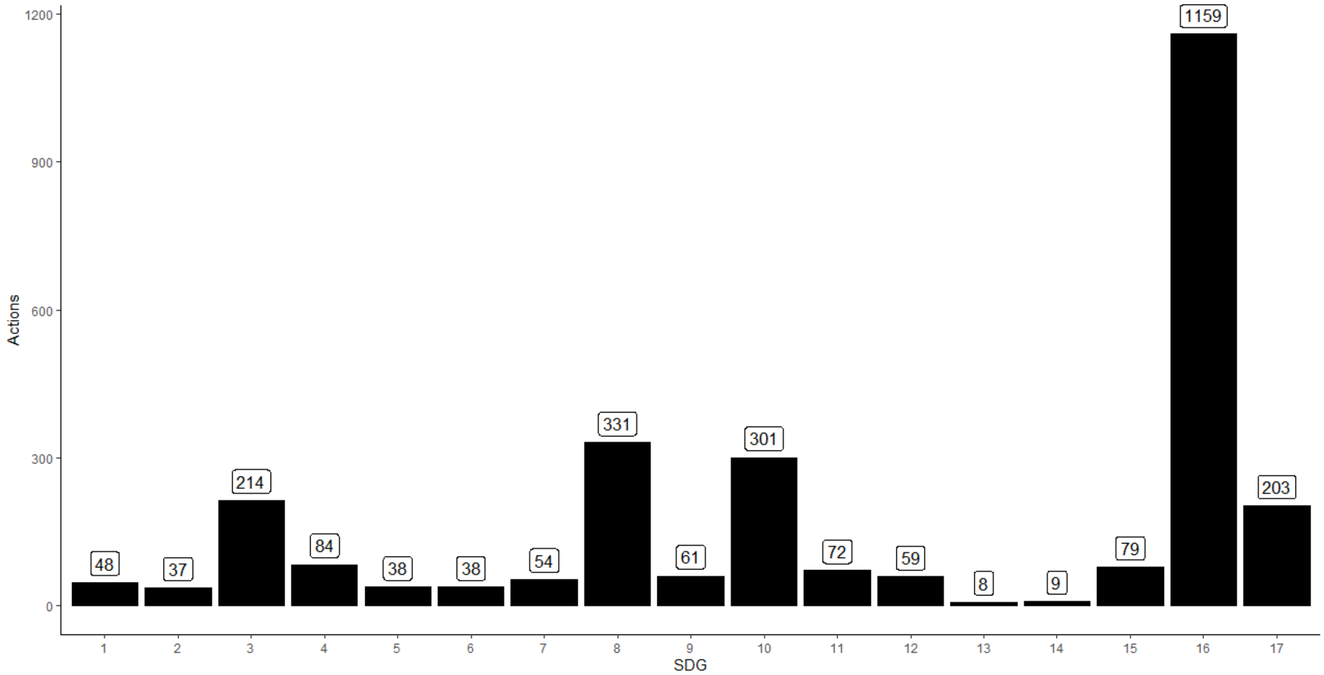

Figure 5. SDGs - Actions count.

Table 6. Example - Text feature (Cont.).

order	text
1	[cats, are, so, cute]
2	[mice, me, scare]
3	[the, cat, hunts, the, mouse]
4	[that, cute, mice, and, cats, playing]
5	[an, army, of, mice, runs, after, the, cat]
6	[the, cat, asks, for, peace, to, the, mouse]
7	[the, cat, is, afraid]
8	[cat, and, mouse, live, in, peace]

Table 7. Example - Text feature (Dictionary).

id	word
1	cats
2	are
3	so
4	cute
5	mice
6	me
	...
27	fear
28	live
29	in

Figure 5 shows SDG frequency count among labeled actions in the databse. It show that SDG 16 - Peace, Justice and Effective Institutions - has much more labels than the others and, in this specific database, it is the record with the greatest potential for machine learning. With some balancing effort, SDG 3, 8 and 10 could be used for training and prediction via Catboost, but the present work will focus only on SDG 16, given the large number of scenarios that will be evaluated. This Sustainable Development Goal is the most complex for classification by the specialists of the court, due to the constitutional character of the Federal Supreme Court. To some extent, almost all processes forwarded to the STF are related to this SDG, which makes the classification more complex. Therefore, it is important to find an algorithm with the potential to automate the classification.

Table 8. Example - Text feature (Final).

order	text
1	[1,2,3,4]
2	[5,6,7]
3	[8,9,10,8,11]
4	[12,13,5,14,1,15]
5	[16,17,18,5,19,20,21,9]
6	[8,9,22,23,24,11]
7	[8,9,25,26,27]
8	[9,14,11,28,29,23]

5. Results and Discussion

Scenarios performance will be compared through accuracy, results are presented in Table 9. 15% of the initial set was set aside for testing. Scenario 5 (with STF actions, SDG keywords count and pos tag step) had the best performance according to the metric accuracy.

It can be seen that in scenarios 2, 4, 6 and 7 there is not

Table 9. Accuracy.

Scenario	Full Text	Text with Post Tag
Scenario 1	0.76	0.79
Scenario 2	0.65	0.66
Scenario 3	0.79	0.75
Scenario 4	0.75	0.75
Scenario 5	0.83	0.85
Scenario 6	0.82	0.83
Scenario 7	0.81	0.81

much difference between using the post tag approach or not. The field of law proved to be the worst performing variable alone. Overall, the accuracies and F1-scores (> 0.7) in almost all scenarios are good, which makes the Catboost algorithm a good candidate for textual classification in this specific legal context. As F1-Score metric combines precision and recall, the possibility of the Catboost algorithm predicting only the majority class, a problem that often occurs in aggressively unbalanced data contexts, is ruled out. This is very good, reassuring model application in production (court routine). Considering actions distributions presented on Figure 2 and 3, we repeated the analysis keeping only cases of the ADI class (STF action) and Administrative Law and other Public Law Matters field. Reducing the universe of analysis and, consequently, the database, the post tag approach proved to be strong, since it reduced the size of the texts without necessarily making them less dense. The accuracy keeping the full texts and the best scenario is 68% and with the post tag approach it exceeds 85%. This is the biggest difference observed in this study and proves the importance of using quality metadata in classification flows. Often, the classification target data has unbalanced classes and little input data. This makes machine learning and deep learning routines difficult. Here, this difficulty was overcome with the use of metadata associated with the model. This is an elegant outlet that is often overlooked in favor of exclusively text-focused models.

The idea of using machine learning and automation in legal contexts, by lawyers and courts, should not be centered on the full replacement of human labor. On the contrary, it is about associating the best features of the machines with the work of specialists, in order to reduce repetitive and manual activities. In terms of business, having a tool capable of using the same metadata that employees and lawyers use to automatically tag texts is very useful for courts and legal techs, due to its speed and efficiency gains, especially when models are associated with dashboard tools such as Shiny Apps [16], Streamlit [17] and Qlik Sense [18].

On the technical side, there are some improvements on the horizon. Text cleaning can also be done by removing unnecessary terms and lemmatization. Regarding the way to vectorize texts, which is natively done by the Catboost algorithm using Bag of Words (BoW), it is also possible to use context tools, such as Doc2Vec. In this case, the vectors generated by Table 8 would all be dense and of the same size, taking advantage of the relationships between words arising from the context of each legal text. Text embeddings produced by transformers architectures [19] are usually quite powerful in representation activities. Its use in legal texts in Portuguese, however, is hampered by the size of the texts, which exceeds the maximum value of tokens in pre-trained transformers. To better serve the actions assessment flows in courts and law offices, new tags can be targeted by automations and machine learning models. The present work has focused on specific UN 2030 Agenda tags, but the method of combining metadata and the Catboost algorithm can be used on any tag.

Author contributions

Lucas José Gonçalves Freitas worked specifically on cleaning the data and adjusting the CatBoost algorithms, while Pamella Sada Dias Edokawa and Euler Rodrigues de Alencar worked on obtaining the data, building the scenarios, choosing the metadata and the legal interpretation of the results. Thaís Carvalho Valadares Rodrigues worked on the model evaluation metrics and on the text review. Ariane Hayana Thomé de Farias made the graphics, tables and took care of the template in LaTeX. All authors participated in the construction of the models and the production of the text.

References

[1] KATZ, D. M.; BOMMARITO, M. J.; BLACKMAN, J. A general approach for predicting the behavior of the supreme court of the united states. *PLoS ONE*, São Francisco, EUA, v. 12, n. 4, p. 485–498, abr. 2017.

[2] MEDVEDEVA, M.; VOLS, M.; WIELING, M. Using machine learning to predict decisions of the european court of human rights. *Artificial Intelligence and Law*, Dordrecht, Holanda, v. 28, p. 237–266, jun. 2020.

[3] HAUSLADEN, C. I.; SCHUBERT, M. H.; ELLIOT, A. Text classification of ideological direction in judicial opinions. *International Review of Law and Economics*, Sevenoaks, Inglaterra, v. 62, p. 1–39, jun. 2020.

[4] RADYGIN, V. Y. et al. Application of text mining technologies in russian language for solving the problems of primary financial monitoring. *Procedia Computer Science*, Amsterdã, Holanda, v. 190, p. 678–683, jan. 2021.

[5] JUNIOR, A. P. C.; CALIXTO, W. P.; CASTRO, C. H. A. Aplicação da inteligência artificial na identificação de conexões pelo fato e tese nas petições iniciais e integração com o sistema de processo eletronico. *Revista Eletrônica do CNJ*, Brasília, v. 4, n. 1, p. 8–18, jan. 2020.

[6] NASCIMENTO, E. G. S.; OLIVEIRA, R. S. Clustering by similarity of brazilian legal documents using natural language processing approaches. In: TANG, N. (Ed.). *Data Clustering*. Rijeka, Croácia: Intechopen, 2022. p. 1–15.

[7] MENEZES, E. J. N.; CLEMENTINO, M. B. M. Using deep learning to predict outcomes of legal appeals better than humans experts: A study with data from brazilian federal courts. *PLoS ONE*, São Francisco, EUA, v. 17, n. 7, p. 1–20, jul. 2022.

[8] ZANUZ, L.; RIGO, S. J. Fostering judiciary applications with new fine-tuned models for legal named entity recognition in portuguese. In: *PROPOR 2022: 15th International Conference on Computational Processing of Portuguese*. Cham, Suíça: Springer, 2022. p. 219–229.

[9] UNITED NATIONS BRAZIL. *Agenda 2030*. 2022. Disponível em: ⟨https://brasil.un.org/pt-br/sdgs⟩. Acesso em: 10 set. 2022.

[10] SUPREMO TRIBUNAL FEDERAL. *Agenda 2030 and STF*. 2022. Disponível em: ⟨https://bit.ly/STFAgendaONU2030⟩. Acesso em: 3 jul. 2022.

[11] SPACY. *spaCy: Industrial-strength natural language processing in python*. 2022. Disponível em: ⟨https://spacy.io/⟩. Acesso em: 3 jul. 2022.

[12] WALSH, M. *Part-of-Speech Tagging for Portuguese*. 2021. Disponível em: ⟨https://melaniewalsh.github.io/Intro-Cultural-Analytics/05-Text-Analysis/Multilingual/Portuguese/03-POS-Keywords-Portuguese.html⟩. Acesso em: 3 jun. 2022.

[13] CATBOOST. *CatBoost documentation*. 2022. Disponível em: ⟨https://catboost.ai/en/docs/⟩. Acesso em: 3 jul. 2022.

[14] UNIVERSIDADE DE BRASÍLIA. *Laboratory of Machine Learning in Finance and Organizations*. 2022. Disponível em: ⟨https://lamfo-unb.github.io/2017/09/27/BaggingVsBoosting/⟩. Acesso em: 3 jul. 2022.

[15] CATBOOST. *CatBoost tutorials*. 2022. Disponível em: ⟨https://github.com/catboost/tutorials⟩. Acesso em: 3 jul. 2022.

[16] CHANG, W. et al. *shiny: Web Application Framework for R*. 2022. Disponível em: ⟨https://shiny.rstudio.com/⟩. Acesso em: 3 jul. 2022.

[17] STREAMLIT. *Streamlit*. 2022. Disponível em: ⟨https://streamlit.io/⟩. Acesso em: 3 jul. 2022.

[18] QLIK SENSE. *Qlik Sense*. 2022. Disponível em: ⟨https://www.qlik.com/⟩. Acesso em: 3 jul. 2022.

[19] SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: *BRACIS 2020: 9th Brazilian Conference on Intelligent Systems*. Cham, Suíça: Springer, 2020. p. 403–417.