

DeepPolicyTracker: Tracking Changes In Environmental Policy In The Brazilian Federal Official Gazette With Deep Learning

Tackling Climate Change with Machine Learning workshop at ICML 2021

Flávio Nakasato Cação ¹

Anna Helena Realí Costa ¹

Natalie Unterstell ²

Liuca Yonaha ²

Taciana Stec ²

Fábio Ishisaki ²

¹ Polytechnic School, University of Sao Paulo, Sao Paulo, Brazil

² Política por Inteiro, Sao Paulo, Brazil

¹ {flavio.cacao, anna.reali}@usp.br

² {natalie, liuca, taciana, fabio}@politicaporinteiro.org

* * *



Motivation

- Brazil has a **key role** in the climate change debate, once it has some of the most important biomes in the world, such as the **Amazon Rainforest** and the **Cerrado**
- However, the country is also one of the **biggest emitters of greenhouse gases** in the world, due to **farming, deforestation** and **forest fires**
- **Regardless** of lobbies or political orientation, **all government legal actions** are published **daily** in the **Federal Official Gazette**

Challenge: how could **machine learning** help to **track hundreds of decrees issued every day** by the authorities?

Our proposal: *DeepPolicyTracker*

- A deep learning model based on BERT in Portuguese, a state-of-the-art pre-trained language model
- **Objectives: tracking** changes in the national **environmental policies** and **classifying government acts**
- To fine-tune the system: **rule-based bot** + **dataset annotated** by **domain experts** *

* The version of the dataset annotated by domain experts and used in this work is available at <https://github.com/nakasato/deepolicytracker>. Its latest version can be consulted directly on the organization's website: <https://www.politicaporinteiro.org/monitor-de-atos-publicos/>

How it works

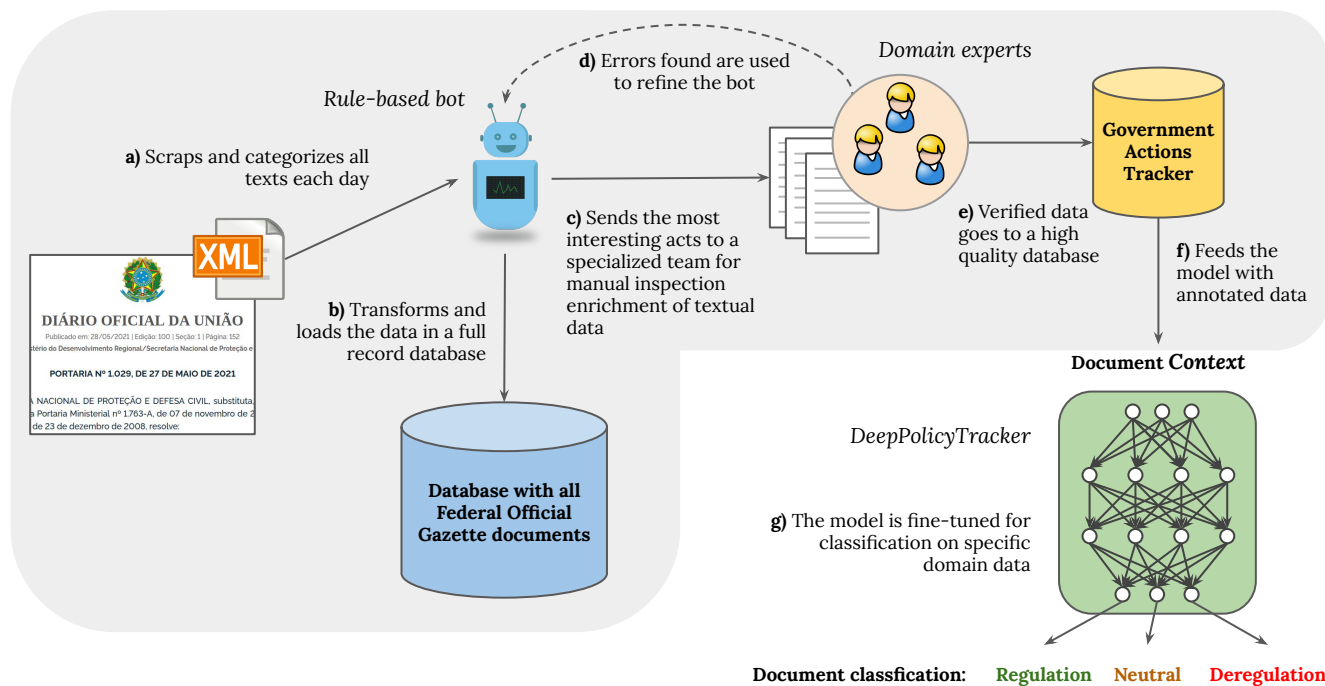


Figure 1. The operational flow of the data pipeline.

How it works

- Domain experts from Política Por Inteiro defined **12 classes** for the **Classification field**, like “Regulation”, “Flexibilization” and “Privatization”
- We **regrouped** the previous classes of the **Classification** variable into **2 training settings**:
 - **With three major classes:**
 - **Regulation:** Regulation, Planning and Response
 - **Neutral:** Neutral, Retreat and Legislation
 - **Deregulation:** Privatization, Deregulation, Flexibilization, Institutional reform, Law consolidation and Revocation
 - **With only two major classes:**
 - **Regulation:** Regulation, Planning, Response, Neutral, Retreat and Legislation
 - **Deregulation:** Privatization, Deregulation, Flexibilization, Institutional reform, Law consolidation and Revocation

Current Results

- Below, Matthews Correlation Coefficient (MCC) and the Accuracy (Acc) results obtained for the two types of rearrangements

Table 1. Summary of the Matthews Correlation Coefficient (MCC) and Accuracy (Acc) results on the test set for the two types of classification, with 2 and 3 classes.

TARGET	MCC	ACC	PROPORTION OF EACH CLASS.
CLASSIFICATION (2 CLASSES)	0.61	0.84	REGULATION (68.7%), DEREGULATION (31.3%)
CLASSIFICATION (3 CLASSES)	0.57	0.74	REGULATION (49.3%), NEUTRAL (19.4%), DEREGULATION (31.3%)

Current Results

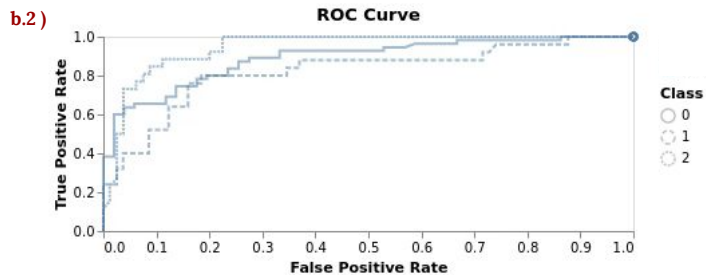
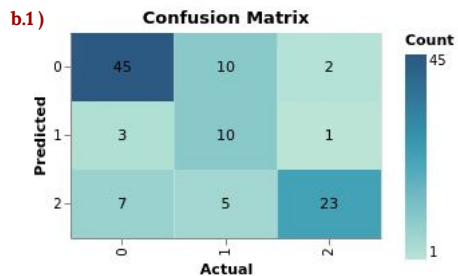
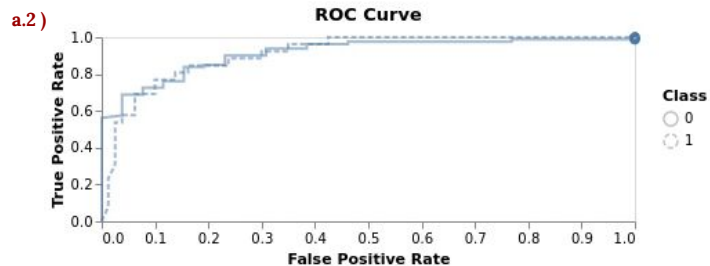
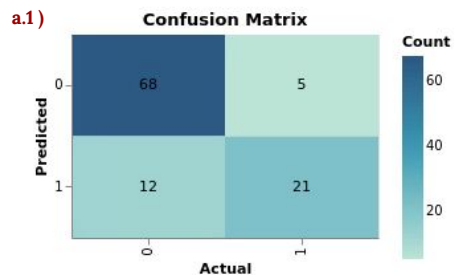


Figure 2. Confusion matrices and ROC curves generated by the models a) with 2 classes (0: *Regulation*, 1: *Deregulation*) and b) with 3 classes (0: *Regulation*, 1: *Neutral*, 2: *Deregulation*).

Conclusion & Future Work

- *DeepPolicyTracker*, a **model in progress** of an end-to-end neural system based on **state-of-the-art NLP** which aims to **track** and **classify** changes in **environmental policies** directly from **texts in official documents**
- Rule-based system + domain experts already deliver **real value**
- Now, it is time to increase the **scalability** and **performance**
- Despite of the extra-small dataset, *DeepPolicyTracker* model is promising
- Focus: **Monitoring** in order to **inform** the civil society
- For the future: improve the stability for a **greater number of classes** and compare to **simpler models too**

Thank you!

Special thanks to:

- Centro de Ciência de Dados da USP (C2D)
- Política Por Inteiro
- University of Sao Paulo
- Polytechnic School