Transfer Learning for Deontic Rule Classification: the Case Study of the GDPR

Davide Liga and Monica Palmirani

March 11, 2023

April 2022

Transfer Learning for Deontic Rule Classification: the Case Study of the GDPR

Davide LIGA [a,b,1], Monica PALMIRANI [b]
[a] *University of Luxembourg*
[b] *Alma Human-AI, University of Bologna*
ORCiD ID: Davide Liga https://orcid.org/0000-0003-1124-0299, Monica Palmirani https://orcid.org/0000-0002-8557-8084

Abstract. This work focuses on the automatic classification of deontic sentences. It presents a novel Machine Learning approach which combines the power of Transfer Learning with the information provided by two famous LegalXML formats. In particular, different BERT-like neural architectures have been fine-tuned on the downstream task of classifying rules from the European General Data Protection Regulation (GDPR) encoded in Akoma Ntoso and LegalRuleML. This work shows that fine-tuned language models can leverage the information provided in LegalXML documents to achieve automatic classification of deontic sentences and rules.

Keywords. Rule classification, Norms, Legal Knowledge Representation, AI&Law

The ability to automatically detect deontic rules directly from natural language sentences is a crucial long-term goal in the field of Artificial Intelligence and Law (AI&Law), and in legal argumentation [1,2]. One of the obstacles of this kind of tasks is the lack of available data designed *ad hoc* for the classification of deontic rules. Since the annotation of this kind of datasets is time-consuming and requires experts of domain, the process of creating datasets to automatically recognize deontic rules can be costly. Another obstacle, related to the first one, is that datasets might be too small to train Machine Learning classifiers, especially when dealing with deep neural architectures.

To tackle these issues, a groundbreaking methodology has recently been employed in AI (and NLP), namely Transfer Learning, an approach where huge pre-trained neural architectures are employed in downstream tasks. In this regard, BERT is one of the most famous examples, used in many downstream tasks even with very small datasets [3,4].

On the one side, this work shows the potential of using LegalXML documents as source of data. On the other side, it exploits the ability of Transfer Learning to have good performances on downstream tasks even when dealing with small datasets. Furthermore, this work tackles the automatic classification of deontic rules directly from natural language, an AI&Law task which has been approached by the community only marginally.

---

[1] Corresponding Author: Davide Liga, davide.liga@uni.lu.
Davide Liga has been founded under the INDIGO project. The project INDIGO is financially supported by the NORFACE Joint Research Programme on Democratic Governance in a Turbulent Age and co-funded by AEI, AKA, DFG and FNR and the European Commission through Horizon 2020 under grant agreement No 822166.

This task consists in classifying single legal sentences or single legal provisions as containing deontic modalities such as Obligations, Prohibitions and Permissions.

1. Methodology

We consider two methodological aspects: the data extraction method (how we retrieved our data) and the classification method (what Machine Learning approach we used). To extract labelled data for the classification of rules and deontic modalities, we combined Akoma Ntoso and LegalRuleML as suggested in [5]. While LegalRuleML is an optimal representation of the legal logical sphere, Akoma Ntoso is an optimal representation of the structure of legal document, including their natural language. This can facilitate the reconstruction of the atomic legal provisions from natural language sentences, especially in those cases where the deontic information is split in different structural portions within the legal source. In this work, these two formats have been used to create a dataset, where atomic legal provisions are taken (and sometimes reconstructed) from Akoma Ntoso, while the logical/deontic classes are extracted from LegalRuleML.

Regarding the classification methodology, we employed Transfer Learning, which consists in the downstream use of language models (i.e., neural architectures that have been pre-trained on a huge amount of data). Importantly, there are two major ways of performing Transfer Learning: a famous approach is to use the pre-trained language model to extract embeddings to represent our data (as described in [4]). Another approach is that of fine-tuning the pre-trained neural architecture on a downstream task. In this work, we used this second approach.

2. Related Works

The first studies which tackled the classification of deontic elements focused on the deontic elements as parts of a wider range of targets [6,7,8] or were strongly based on symbolic approaches of Artificial Intelligence [9,10]. Perhaps, the first studies which mainly focused on the deontic sphere are [11,12], which are also among the first which employed sub-symbolic methods such as Bi-LSTM and self-attention in the field of AI&Law.

Since the publication of BERT [3], a growing number of studies employed Transfer Learning methods. To the best of our knowledge, the first study which employed BERT for the classification of deontic sentences is [13]. While [12] focused on just prohibitions and obligations, [13] also focused on permissions, achieving an average precision and recall of 90% and 89.66% respectively. Another recent work is [14], which used four pre-trained architectures (BERT, DistilBERT, RoBERTa, and ALBERT) but focused just on the binary detection duties vs non-duties.

Also our work presents a Transfer Learning approach based on BERT (and other similar models), which leverage the symbolic information of LegalXML formats (see also [5]) exploiting the sub-symbolic power provided by different pre-trained language models. The novelty and the power of Transfer Learning methodologies jointly with the combined use of Akoma Ntoso and LegalRuleML are two major contributions of our study, along with the design of the experimental settings in 4 different classification scenarios: (1) Rule vs Non-rule, (2) Deontic vs Non-deontic, (3) Obligation vs Permission vs

None, (4) Obligation vs Permission vs Constitutive Rule vs None. Another point which is worth mentioning is that LegalXML formats such as Akoma Ntoso and LegalRuleML are documents which are written by legal experts, providing the machine learning algorithm with high quality data.

3. Data

The data used in this study consists of 707 atomic normative provisions[2] extracted from the European General Data Protection Regulation (GDPR). To extrapolate this dataset, we used the *DAta Protection REgulation COmpliance* (DAPRECO) Knowledge Base[15], which is the LegalRuleML representation of the GDPR and the the biggest knowledge base in LegalRuleML [16], as well as the biggest knowledge base formalized in Input/Output Logic [17]. The current version of the DAPRECO[3] includes 966 formulæ in reified Input/Output logic: 271 obligations, 76 permissions, and 619 constitutive rules. As explained in [15], the number of constitutive rules is much higher than permissions and obligations because constitutive rules are needed to trigger special inferences for the modelled rules. This means that constitutive rules are an indicator of the existence of a rule, without giving information about deontic modalities.

Importantly, DAPRECO also contains the connections between each formula and the corresponding structural element (paragraphs, point, etc) in the Akoma Ntoso representation of the GDPR[4]. In other words, using a LegalRuleML knowledge base like DAPRECO and the corresponding Akoma Ntoso representation, it is possible to connect the logical-deontic sphere of legal documents (in this case the 966 Input/Output formulæ provided by DAPRECO) to the natural language statements in the legal text (provided by the Akoma Ntoso representation of the GDPR).

Importantly, this combination of Akoma Ntoso and LegalRuleML facilitate also the reconstruction of the exact target in terms of natural language. For example, many obligations of legal texts are split into lists, and Akoma Ntoso is useful to reconstruct those pieces of natural language into a unique sentence.

4. Experiment settings and results

At the end of the process of extraction, we achieved a total of 707 labelled provisions, which have been reconstructed whenever they were split into lists (thanks to the structural information provided by Akoma Ntoso). The labels of these sentences are the same as those provided by DAPRECO with the addition of a "none" category. We abbreviated "obligationRule", "permissionRule", "constitutiveRule" in "obligation", "permission" and "constitutive" respectively.

The class "obligation" is referred to those sentences which have at least one obligation rule in their related formulæ. The class "permission" is referred to those sentences

[2]These provisions belong to the body of the GDPR (preamble and conclusions were excluded), and are generally paragraphs or list points, which may sometimes consist of multiple sentences.

[3]The DAPRECO knowledge base can be freely downloaded from: `https://github.com/dapreco/daprecokb/blob/master/gdpr/rioKB_GDPR.xml`.

[4]The Akoma Ntoso representation of the GDPR can be currently accessed from `https://github.com/guerret/lu.uni.dapreco.parser/blob/master/resources/akn-act-gdpr-full.xml`.

Table 1. Number of instances per class per scenario.

	Classes	Instances
Scenario 1	rule	260
	non-rule	447
Scenario 2	deontic	204
	non-deontic	503
Scenario 3	obligation	156
	permission	44
	none	503
Scenario 4	obligation	156
	permission	44
	constitutive	56
	none	447

Table 2. Results for the two stratified baselines applied to scenario 1, 2, 3 and 4. Within the brackets, the number of instances is reported. P = precision; R = recall, F1 = F1-score; Acc = Accuracy; Mcr = Macro F1.

Baseline Scenario 1

	P	R	F1	Acc	Mcr
rule(39)	.37	.33	.35	.55	.50
non-rule(67)	.63	.67	.65		

Baseline Scenario 2

	P	R	F1	Acc	Mcr
deontic(30)	.23	.23	.23	.57	.47
non-deontic(76)	.70	.70	.70		

Baseline Scenario 3

	P	R	F1	Acc	Mcr
obligation(24)	.24	.17	.20	.60	.40
permission(6)	.20	.33	.25		
none(75)	.73	.76	.75		

Baseline Scenario 4

	P	R	F1	Acc	Mcr
obligation(23)	.16	.13	.14	.44	.23
permission(7)	.20	.14	.17		
constitutive(8)	.00	.00	.00		
none(67)	.58	.63	.60		

which have at least one permission rule in their related formulæ. The class "constitutive" is referred to those sentences which just constitutive rules in their related formulæ. The class "none" is referred to all sentences which have no rule at all. These labels allowed 4 different experimental settings, as shown in Table 1.

Scenario 1 is a binary classification and aims at discriminating between rule and non-rule instances. In this scenario, all labels other than "none" are considered rule, while "non-rule" is just an alias for "none". Scenario 2 focus on a binary classification between deontic instances (i.e., any sentence labelled as either "obligation" or "permission") and non-deontic instances (i.e., all instances which are labelled neither as "obligation" nor as "permission"). Scenario 3 is a multiclassification which considers the classes "obligation", "permission" and "none" (with "constitutive" considered as part of the latter). Scenario 4 is a multiclassification which considers the classes "obligation", "permission", "constitutive" and "none". For the multi-classifications (i.e. Scenario 3 and 4) four statements have been removed, since the classes "obligation" and "permission" overlapped.

To assess the non-triviality of our experiments, we employed different kinds of baseline, showing their difficulty in performing each classification task. In Table 2, we reported the results of a "stratified" baseline, which was the better performing among the baseline methods[5]. As can be seen from Table 2, we applied the "stratified" baseline on all the scenarios achieving quite low performances on all of them.

[5] The so-called "stratified" baselines can reach higher scores because they reflect the class distribution.

Table 3. Results of the 4 scenarios. P = precision; R = recall, F1 = F1-score, S/T = Support/Total ratio.

Scen.	Classes	BERT			DistilBERT			LegalBERT			S/T
		P	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	
1	rule	.74	.95	.83	.77	.95	.85	.75	.77	.76	39/260
	non-rule	.96	.81	.88	.97	.84	.90	.87	.85	.86	68/447
		Accuracy .86 Macro avg .86 Weight. avg .86			Accuracy .88 Macro avg .87 Weight. avg .88			Accuracy .82 Macro avg .81 Weight. avg .82			Total: 107/707
2	deontic	.74	.90	.81	.82	.90	.86	.80	.77	.79	31/200
	non-deontic	.96	.87	.91	.96	.92	.94	.91	.92	.92	76/507
		Accuracy .88 Macro avg .86 Weight. avg .88			Accuracy .92 Macro avg .90 Weight. avg .92			Accuracy .88 Macro avg .85 Weight. avg .88			Total: 107/707
3	obligationRule	.74	.83	.78	.74	.83	.78	.63	.92	.75	24/156
	permissionRule	.50	.83	.62	.36	.67	.47	.56	.83	.67	6/44
	none	.97	.88	.92	.94	.84	.89	1.0	.82	.90	76/503
		Accuracy .87 Macro avg .78 Weight. avg .88			Accuracy .83 Macro avg .71 Weight. avg .84			Accuracy .84 Macro avg .77 Weight. avg .85			Total: 106/703
4	obligationRule	.70	.79	.75	.80	.83	.82	.84	.67	.74	24/156
	permissionRule	.60	.50	.55	.40	.67	.50	.17	.67	.28	6/44
	constitutiveRule	.36	1.0	.53	.47	.89	.62	.89	.89	.89	9/56
	none	1.0	.73	.84	.94	.76	.84	.96	.79	.87	67/447
		Accuracy .75 Macro avg .67 Weight. avg .78			Accuracy .78 Macro avg .69 Weight. avg .80			Accuracy .76 Macro avg .69 Weight. avg .81			Total: 106/703

As far as the experimental settings are concerned, the dataset was divided into 70% for the training phase, 15% for the test and 15% for the validation; and for all instances, a max length of 30 was applied.

Regarding the Transfer Learning architecture, 3 pre-trained language model have been fine-tuned, namely BERT [3], DistilBERT [18], and the LegalBert trained on EurLex [19]. These three neural architectures were fine-tuned by adding two linear layers with a ReLu activation function and with a dropout of 0.2 after each activation, and a final output layer was added for the classification, through a softmax activation function. The fine-tuning process of these 3 neural architectures was performed in 10 epochs (learning rate: 1e-3; batch size: 32).

The final results on the validation set are reported in Table 3, where it can be seen that DistilBERT outperforms the other classifiers in the binary classifications, with an average score reaching .88 in the first scenario and .92 in the second one.

The results for the third and fourth scenarios are less straightforward and show that BERT slightly outperforms other classifiers in the third scenario, while LegalBERT outperformed the other models in the fourth scenario. The main problem for the multiclassifications, is the class unbalance and the restricted amount of instances for some classes. In spite of this, scores are encouraging, especially considering the small amount of data.

5. Conclusions

The contribution of this work is showing how Transfer Learning methods can leverage the information provided in LegalXML to train classifiers capable of automatically classifying deontic sentences and rules. In the future, we would like to create a stronger connection with the ontological sphere by using PrOnto [20], strengthening this hybrid AI approach, which combines symbolic knowledge with sub-symbolic methods.

In this work, we were not targeting in the internal elements of the logical formulæ, we just addressed the ontological classes of each rule. However, in the future we want to create classifiers that directly address the internal components of each rule, trying to find a match between portions of natural language and portions of rules. In general, the ability to connect each internal component (or at least some) of the deontic formulæ contained in DAPRECO directly to the portion of natural language where the component is communicated or expressed is a crucial future direction, and an important step towards the long-term goal of filling the gap between natural language and the logical-inferential sphere, which would generate a more reliable and explainable Artificial Intelligence.

References

[1] Ashley KD. Artificial intelligence and legal analytics: new tools for law practice in the digital age. 2017.
[2] Wyner A, Peters W. On rule extraction from regulations. 2011:113-22.
[3] Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.
[4] Liga D, Palmirani M; 20. Transfer Learning with Sentence Embeddings for Argumentative Evidence Classification. Proceedings of the 20th Workshop on Computational Models of Natural Argument. 2020.
[5] Liga D, Palmirani M. Deontic Sentence Classification Using Tree Kernel Classifiers. 2023:54–73.
[6] Kiyavitskaya N, Zeni N, Breaux TD, Antón AI, Cordy JR, Mich L, et al.; Springer. Automating the extraction of rights and obligations for regulatory compliance. 2008:154-68.
[7] Waltl B, Muhr J, Glaser I, Bonczek G, Scepankova E, Matthes F. Classifying Legal Norms with Active Machine Learning. 2017:11-20.
[8] Gao X, Singh MP. Extracting normative relationships from business contracts. 2014:101-8.
[9] Dragoni M, Villata S, Rizzi W, Governatori G. Combining Natural Language Processing Approaches for Rule Extraction from Legal Documents. 2018:287–300.
[10] de Maat E, Winkels R. Automated Classification of Norms in Sources of Law. 2010:170–191.
[11] Neill JO, Buitelaar P, Robin C, Brien LO. Classifying sentential modality in legal language: a use case in financial regulations, acts and directives. 2017:159-68.
[12] Chalkidis I, Androutsopoulos I, Michos A. Obligation and prohibition extraction using hierarchical rnns. arXiv preprint arXiv:180503871. 2018.
[13] Joshi V, Anish PR, Ghaisas S. Domain adaptation for an automated classification of deontic modalities in software engineering contracts. 2021:1275-80.
[14] Shaghaghian S, Feng LY, Jafarpour B, Pogrebnyakov N; IEEE. Customizing Contextualized Language Models for Legal Document Reviews. 2020:2139-48.
[15] Robaldo L, Bartolini C, Palmirani M, Rossi A, Martoni M, Lenzini G. Formalizing GDPR provisions in reified I/O logic: the DAPRECO knowledge base. Journal of Logic, Language and Information. 2020;29(4):401-49.
[16] Athan T, Boley H, Governatori G, Palmirani M, Paschke A, Wyner A. Oasis legalruleml. In: Proceedings of the Fourteenth International Conference on AI and Law; 2013. p. 3-12.
[17] Makinson D, Van Der Torre L. Input/output logics. Journal of philosophical logic. 2000;29(4):383-408.
[18] Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:191001108. 2019.
[19] Chalkidis I, Fergadiotis M, Malakasiotis P, Aletras N, Androutsopoulos I. LEGAL-BERT: The muppets straight out of law school. arXiv preprint arXiv:201002559. 2020.
[20] Palmirani M, Martoni M, Rossi A, Bartolini C, Robaldo L. Pronto: Privacy ontology for legal compliance. 2018:142-51.