Problem Overview

- Although legal cases are usually represented in textual form, computational analysis has not been widely implemented in legal judgment prediction.
- Methods of natural language processing (NLP) based on neural-network architectures have shown impressive accuracy in predicting the outcomes of legal cases solely based on textual facts provided by the claimants.
- We built transformer-based neural networks, achieving state-of-the-art results on binary and multi-label classification problems in the field of legal judgment prediction, uncovering the potential of NLP to serve as an aid for judges while helping citizens access the fairness of judgments.
- As part of our work, we propose novel hierarchical network architectures in a multi-task setting showing great promise in both performance and explainability to generate decision rationales based on case facts.

Methods

- In approaching our problem, we trained models on two downstream tasks of human rights article violation - binary (any article) and multi-label classification (specific article) - in three increasingly complex steps:
  1. We used pre-trained versions of three large language models from Hugging Face, i.e., BERT, RoBERTa, and LEGAL-BERT and fine-tuned them on our dataset, performing a hyperparameter grid search on data subsets. We only used the first 512 tokens of every case due to BERT-based token limits.
  2. Next, we built custom hierarchical models using any of the above models as a base. Each paragraph was fed through the base model and the resulting embeddings were combined into a case embedding via multi-head attention or transformer layers, as per our specification in Fig. 1.
  3. Finally, we introduced Automated Legal Expert Arbitrator (ALExa), a multi-task hierarchical language model with self-learning loss weights \[ \text{using attention forcing} \] to learn legal judgement rationales (loss weighting function in Equation 1). ALEXa uses BERT as the base model and the Chalkidiki 2021 dataset “rationales for attention forcing, where available for each case (Fig. 2).

Analysis

- Two key issues in legal judgment prediction identified by Chalkidis et al. (2019) are that most systems have severe limitations in “processing long documents” and provide “no justification for their predictions.”
- By building trainable hierarchical models which first embed paragraph meaning and then use multi-head attention or transformer layers to produce a final case embedding, we successfully process longer texts.
- Justifications in the legal domain are most useful on a fact (paragraph) level as opposed to token-level attention scores. By introducing ALEXa, we go beyond paragraph attention to make legal fact selection an explicit component of the training procedure to improve the state-of-the-art.
- Given our limited resources, through grid search on data subsets we found that processing 48 paragraphs with 224 word tokens each using a learning rate of 2e-5 worked best. This can likely still be improved.
- We also conducted a thorough qualitative analysis of ALEXa, showing that it can effectively select the relevant paragraphs in legal cases (Fig. 3).

Conclusions & Limitations

- Our state-of-the-art results for both the binary and multi-label classification tasks underscore the potential of domain-pre-trained and hierarchical language models in legal judgment prediction.
- Given the limited data, additional data would be useful to us, and we are confident we can further improve our results.
- Multi-label hierarchical model performance remains a limitation.

References