

PREVIOUS OPINIONS IS ALL YOU NEED

Legal opinions retrieval system.

INTRODUCTION

Over the past few years, the development of NLP practical solutions has been very dynamic and this advancement led to massive improvements in the quality of AI models. This work focused on the niche domain of the legal field and legal technology applications. Every lawyer knows the pain of having to spend hours drafting and writing legal opinions, briefs, and other legal documents.

Legal opinions are expensive to produce and time-consuming to deliver. Our solution, an AI-powered virtual assistant, offers attorneys a new way to deliver legal opinions in just a few seconds. With our document automation and machine learning algorithms, we can read past documents, data mine existing information, and use retrievals to produce legal opinion letters in a fraction of the time of traditional methods. We present a retrieval system that facilitates and accelerates the work of lawyers. The task of the transformer-based model is to match one or more legal opinions from the database to a new query from the client, this allows the person providing legal advice to automate part of their work.

Can an employer reduce an employee's salary?

An employer may reduce an employee's salary if the following conditions are met 1) the employee ...

The salary reduction should be agreed with ...

According to the recently introduced act, if an employer wants to reduce the employee's working time or ...

DATA

The corpus we had at our disposal contained about 10,000 legal cases. It was crucial that each record contained a legal question and answer in the text. Two main elements in preprocessing can be distinguished. Specifically, the first is to remove redundant texts from legal opinions such as encouraging contact with the law firm, e.g. *More information can be found on our fb*. The second important component is handling with non-informative questions like *What possibilities does the client have?*. We developed an algorithm for identifying unspecified questions and then concatenated the extracted context to such queries.

QUESTION

PREPROCESSING ENCODER

- identification of non-informative questions
- extraction of key information from text

LEGAL ANSWER

PREPROCESSING ENCODER

- deletion redundant fragment of the answer

SIMILARITY SCORE

EXPERIMENTS

The research began with testing three models: **LaBSE** [1], **Herbert** [2] and a model specific to the QA problem - **multi-QA** (6 layer version of MiniLM [3] finetuned on 215M QA pairs) in a zero-shot learning configuration. Surprisingly high results were achieved by the LaBSE model in this way (see Table 1). The probability of finding an appropriate legal opinion is more than 40% cases when the model has not seen the data before. However, after model finetuning, we noticed that Herbert gains the highest score and finds semantic relations best in legal texts.

Top N	Zero-shot			Finetuned		
	LaBSE	Herbert	multi-qa	LaBSE	Herbert	multi-qa
1	0.413	0.011	0.217	0.459	0.604	0.423
2	0.520	0.013	0.295	0.488	0.730	0.518
3	0.564	0.015	0.323	0.643	0.792	0.596
5	0.644	0.020	0.384	0.710	0.844	0.665
7	0.703	0.028	0.426	0.763	0.865	0.714
10	0.737	0.032	0.451	0.811	0.888	0.752

Table 1

Moreover, numerous analyses have been conducted, among others, adding negative examples during training or an attempt to match the question with tokenized sentences from legal opinion. Nevertheless, classic training performed best, i.e. comparing the similarity between the full text of the question and the answer. In the next chapter, only the Herbert model will be considered due to its high performance.

RESULTS

The final results are shown in the plot below. To experimentally test our method, we used two sets: 1) **Test data** - the answer search space consisted only of examples that the model had not seen before, 2) **Full data** - seeking answers from all the legal opinions collected so far. The assessment methods on the above datasets accepted only one correct answer, whereas in practical application many legal opinions may prove helpful in reaching a new verdict. Therefore, specialists were given an additional sample of the data to check whether the model predictions would be useful. Based on the lawyer's expertise, we estimated the retriever's realistic performance (light blue on the chart). In 80% of cases, a helpful opinion already appears among the first two model predictions.

Retriever results



CONCLUSION

The usefulness of algorithms help has been appreciated in many aspects of our lives, and AI in our workplaces is becoming increasingly ubiquitous. This study shows that it is possible to create a simple tool that will invaluablely support and optimize the work of lawyers. In creating this system, we were guided by its usability and this condition has been met, as the model achieves high results even when seeking answers among all opinions.

BIBLIOGRAPHY

- [1] Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan and Wei Wang. 2020. **Language agnostic BERT sentence embedding**. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. pp. 878-891. Dublin, Ireland.
- [2] Robert Mroczkowski, Piotr Rybak, Alina Wróblewska and Ireneusz Gawlik. 2021. **HerBERT: Efficiently pretrained transformer-based language model for Polish**. In *Proceedings of the 8th Workshop on BaltoSlavic Natural Language Processing*, pp. 1-10. Kyiv, Ukraine. Association for Computational Linguistics.
- [3] Wanhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang and Ming Zhou. 2020. **MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformer**. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*. pp. 13. Vancouver, Canada.