

TREO

Technology, Research, Education, Opinion

Legal NLP in English to Analyze the Independent Contractor-Employee Classification

Development of a mobile app to support legal deliberations

Lutfus Sayeed, San Francisco State University, lsayeed@sfsu.edu;

Md. Alimoor Reza, Drake University, md.reza@drake.edu

The present project will develop a mobile application to support legal professionals' analyses of labor court judgments to distinguish between independent contractors and employees. To accomplish this objective, the current initiative will implement a Natural Language Processing (NLP) approach to identify the patterns in past rulings and opinions by labor court judges. The mobile application will provide recommendations to legal professionals engaged in litigations related to the appropriate classification issue. The use of NLP will add to the body of research that attempted to delineate the judgment patterns in the classification decision.

Denton et al. [1] employed multi-layer perceptron (MLP), statistical modeling, and expert analyses to develop an analytical framework for the binary classification task. The present effort will extend the previous research in this area by employing recent advances in NLP. Transformer-based models have shown great potential in solving various NLP tasks with impressive performance. This accomplishment can largely be attributed to pretraining these models using large-scale text corpus collected from the internet and then further fine-tuning the model on task-specific smaller datasets at hand [6,7]. Unlike general-purpose NLP tasks that have large-scale text corpus, there are fewer datasets for legal NLP tasks. Moreover, much prior work is based on non-English legal datasets [4, 5] or dataset that is proprietary [3]. The current project will improve previous investigations in this area — labor court judgments to distinguish between independent contractors and employees — by leveraging the availability of a significantly larger amount of data in the English language from an online repository of US legal documents such as LexisNexis [8] and enhancement of computing capabilities to process the readily accessible data. Utilizing our newly introduced dataset, this project will explore transformer-based models (e.g., BERT and its variants) to solve a binary legal judgment prediction task, i.e., predicting the binary outcome (independent contractor vs. employee) of a labor court case given a text describing the details of the case. As prior work reported mixed reactions on domain-specific pretraining [2], to assess the usefulness of additional domain-specific pretraining on the proposed legal judgment prediction task, the project will further investigate the effect of domain-specific pretraining using the proposed dataset. Finally, the project plans to publicly release the dataset and code to advance the research in legal language processing.

References

1. Denton, J.W., Sayeed, L., Moorman, A.H., and Perkins, N.D. "Using Neural Network to Classify Employees for Tax Purposes," *Accounting, Management and Information Technologies*, Vol. 5, No. 2, 1995, pp. 123-138.

Remaining references will be available upon request