COSE474-2024F: Final Project Proposal
"A model for predicting Korean court rulings using RoBERTa"

Moogeun Park

1. Introduction

The website of the Korean courts contains records of various judgments from all civil, criminal, and other cases held so far, including the details of each judgment, the statements of the defendant and plaintiff, and the court's decisions. A single judgment includes a wide range of data, and there are previous cases and appellate court decisions, along with relevant legal articles as the basis for these decisions. The goal is to create an AI model that predicts the plaintiff or defendant's success by utilizing deep learning trained on this data.

2. Problem definition & chanllenges

The main goal is to predict whether the plaintiff or defendant will win, given basic information about the case, facts related to the case, and the arguments of the plaintiff and defendant.

3. Related Works

Research on predicting related judgments and reasoning the judge's cognitive process, particularly using NLP in deep learning, has already been actively conducted. This study aims to design a deep learning model that delivers actual judgments by utilizing the latest NLP techniques.

4. Datasets

First, the case type and subtypes are categorized, and the arguments of the plaintiff and defendant, as well as case-related facts, are gathered from court websites and publicly available judgments. After compiling this data, a label is assigned to each dataset indicating whether the plaintiff or defendant won.

5. State-of-the-art methods and baselines

Although no related benchmarks exist, the focus will be on maximizing the accuracy of judgment predictions.

References

Park Ye Chan, L. J. Deep learning for predicting korean court judgments based on judges' cognitive reasoning. *Journal of AI Humanities*, 13:73–107, 2023.

Sung Won Kim, G. R. P. Deep learning based semantic similarity for korean legal field. *KTSDE*, 11(2):93–100, 2022.