“AI + Law”
An Intelligent Legal Aid System

Weipeng Cao, PhD candidate
2018.10.15
01  Why do we need the system? -Motivation
02  How to use it? -An example
03  What is the technic in it? -NLP, Deep Learning
Motivation

- step1
- step2
- step3
- step4

Inefficient!
Motivation

- The similar cases and corresponding results
- The related legal provisions

Data driven + Human knowledge
An Example

Word segmentation and keywords extraction

Intelligent legal aid system

“刘某与广州市越秀区教育局的人事争议”
Personnel disputes between Liu and the Education Bureau of Yuexiu District, Guangzhou

The related legal provisions
The similar cases

System brain:
Deep leaning
Knowledge base

Case matrix:
64*64

“学校”一词的向量表示
Vector representation of the Chinese word "school"
An Example

“刘某与广州市越秀区教育局的人事争议”
Personnel disputes between Liu and the Education Bureau of Yuexiu District, Guangzhou

The decision issued by the Supreme Court:

- 《劳动合同法》第十条，《劳动合同法》第十四条，《劳动合同法》第四十八条
- 《中华人民共和国民事诉讼法》第十三条，《中华人民共和国民事诉讼法》第一百七十条

The output of our intelligent legal aid system:

- 《劳动合同法》第十条，《劳动合同法》第十四条，《劳动合同法》第四十八条
- 《中华人民共和国民事诉讼法》第十三条，《中华人民共和国民事诉讼法》第一百八十条

Performance evaluation:

- Precision: $4/5 = 0.8$
- Recall: $4/5 = 0.8$
- $F1: 2 \times \text{Precision} \times \text{Recall}/(\text{Precision} + \text{Recall}) = 0.8$
"AI + Law" project: Intelligent legal aid system

Training data:
- Massive legal judgment documents
  (each document = fact description + corresponding judgment result)
- Chinese laws and regulations documents

Model learning:
- Extract the top 64 key words and the corresponding legal provisions from each document
- Transfer each key word to a 64-dimension vector;
  → each document: 64*64 matrix
- Input: Keywords matrix of a legal case;
  Output: legal provisions

Natural language processing (NLP), such as TF-IDF, NER, etc.

Multilabel classification problem
Future works

01 Use some advanced algorithms (e.g., Deep residual networks, Bayesian network) to improve the performance of the “Intelligent legal aid system”.

02 Apply random neural network algorithms including my earlier published algorithms to improve the efficiency of the system.

03 Add more functions to the system, such as automatically generate a indictment for users.
Thank you!