ISSN: 2688-8645

Law-net: A New Method for Legal Text Mining

Xiao Li[1,*] and Jing Li[1]

[1]School of Information Management for Law, China University of Political Science and Law, Beijing, 102249, China

Abstract

Unstructured legal texts, including articles, court records, publications and similar cases, are used to store and analyze the results and find a lot of knowledge. A major challenge for data mining research is to develop effective methods to transform legal texts into well structured forms, so that big data and artificial intelligence systems can generate and verify models on this basis. In this paper, we will give a comprehensive overview of the research and development in this direction. This paper proposes a new method of legal text mining to deal with unstructured legal text, so as to find a lot of knowledge. Specifically, we focus on scalable, effective, and weakly supervised methods, working in legal texts (e.g., articles of law, judicial documents).

Keywords

prediction of judgment; Semantic region recognition; Criminal and civil cases; Legal text mining; Data mining.

1. Introduction

With the continuous improvement of the national legal system and the improvement of laws and regulations, people's legal consciousness is becoming stronger and stronger [1]. At present, as more and more legal texts are open, a large number of texts have been accumulated for the excavation of criminal trial. Therefore, through the collection of texts in the field of law and crime, a corpus in the field of justice has been constructed, And the use of natural language processing technology for mining, is of great significance, we call it legal intelligence, legal intelligence includes the following several application points.

First, the judgment prediction of case documents: according to the description of the case, the final judgment result is predicted.

Secondly, the decision prediction of topological structure prediction: find the dependency between subtasks through the judge's decision logic.

Third, the introduction of discriminative attribute of charge prediction, including the corresponding prediction of low-frequency charge and confused charge: by introducing explicit attribute, low-frequency charge can be judged based on attribute, and confused charge can be distinguished; In addition, multi task learning and attention mechanism can be used to train the attribute based crime prediction model.

Fourthly, the prediction of cause of action based on hierarchical structure: through the hierarchical structure of criminal cause of action (accusation) and civil cause of action, combined with the text information of cause of action itself, using sequence prediction and attention mechanism based on the name of cause of action to train the corresponding model.

Finally, the judgment prediction based on Legal reading comprehension: because the judgment results in civil cases need to be combined with the specific demands of the plaintiff, we can establish a reading comprehension mechanism to simulate the reading comprehension behavior of "people looking for answers with questions" to predict the judgment.

ISSN: 2688-8645

2. Method Overview

This method consists of the following two parts. First of all, the first part is the construction of legal judgment documents, the second is the construction of criminal cases based on big data. Lastly, the third part is case semantic region recognition based on the judicial domain.

2.1. Legal Judgment Document Construction

The purpose of interpretation and reasoning of judgment documents is to improve the acceptability of judgment and realize the organic unity of legal effect and social effect by clarifying the formation process and justification of judgment conclusion [2]. Its main values are to enhance the fairness and transparency of adjudication, standardize the exercise of judicial power, enhance judicial credibility and authority, give full play to the role of adjudication in determining disputes and guiding values, promote socialist core values, strive to make the people feel fairness and justice in every judicial case, and earnestly safeguard the legitimate rights and interests of litigants, Promote social harmony and stability.

If the criminal defendant and his defender apply for the exclusion of illegal evidence, the judgment document shall state whether to investigate the legality of evidence collection, whether to exclude the evidence and the reasons. If a civil or administrative case involves a dispute over the distribution of the burden of proof or the standard of proof, the judgment document shall state the reasons.

The judgment documents should be interpreted and reasoned according to the facts or facts in dispute, based on the court trial, cross examination, court debate and court investigation and verification of evidence. When finding facts based on indirect evidence, we should argue about whether there is a confirmatory relationship between indirect evidence and whether a complete proof system can be formed. When the presumption method is used to determine the fact, the reasons for the start of presumption, the facts and reasons for refutation should be explained, and the formation process of the judgment should be explained.

If there is no dispute on the application of the law and the legal meaning does not need to be clarified, the judgment documents should focus on the content and scale of the judgment. If the parties to the lawsuit have disputes over the application of the law in the case or the legal meaning needs to be clarified, the judge shall respond to the focus of the legal dispute item by item and explain the reasons. If there is a concurrence or conflict of legal norms in the application of law, the judgment document shall state the reasons for the choice. If there is no clear legal provision as the direct basis for adjudication in civil cases, the judge should first look for the most similar legal provision to make a judgment; If there is no most similar legal provisions, judges can make judgments according to customs, legal principles, legislative purposes, etc., and reasonably use legal methods to fully demonstrate and reason the basis of judgment. When a judge exercises his discretion to deal with a case, he should adhere to the principles of legality, rationality, impartiality and prudence, fully demonstrate the basis for the use of discretion, and clarify the relevant factors considered in the discretion.

The interpretation and reasoning should be strengthened in the following cases: difficult and complicated cases; The cases in which the parties to the lawsuit are more controversial; Cases with high social concern and great influence; Cases of acquittal, penalty below the legal penalty or death penalty; In administrative litigation, the normative documents on which the administrative action is based are examined together; The case of changing administrative act by judgment; New types of cases or cases that may become guiding cases; Protest cases; A case in which the judgment of the second instance was changed or remanded for retrial; Retrial of cases; Retrial cases; Other cases that need to strengthen reasoning.

The following case judgment documents can simplify the interpretation and reasoning: cases that apply to civil summary procedure and small claims procedure; Cases to be tried by special

ISSN: 2688-8645

civil procedure, supervision procedure and public summons procedure; Cases to be tried by criminal quick adjudication procedure and summary procedure; Minor criminal cases in which the parties reach a settlement agreement; Cases to be tried by administrative summary procedure; The common procedure should be applied to the trial of the case, but the dispute between the parties is not big; Other cases suitable for simplifying reasoning.

The judgment document of the second instance or retrial shall strengthen the interpretation and reasoning of the appeal, protest, and application for retrial. If the facts identified in the judgment documents of the second instance or the retrial are different from those in the first instance or the first instance, or if it is considered that the facts identified in the first instance or the first instance are not clear or that the application of law is wrong, it is necessary to make targeted reasoning on the basis of finding out the facts and correcting the errors in the application of law; In the first instance or the first instance, the reasons have been explained in detail, and there is no dispute or new evidence or new reasons among the parties to the lawsuit, so the interpretation and reasoning can be simplified.

2.2. Criminal Cases Construction Based on Big Data

There are a large number of crimes in big cities, which are much higher than those in other regions [3]. Therefore, the police departments in big cities are extremely eager to better understand and predict criminal incidents, so as to prevent and avoid the harm brought by criminal incidents. Chicago police try to solve this problem through data analysis.

Chicago police first built its analysis data set based on the city website of Chicago, where all the criminal records of Chicago in the past decade are stored. These data record the information of crime in detail, including the detailed criminal record, time and place of crime, processing information and punishment information, etc., and based on the needs of analysis, the corresponding collation is carried out for further analysis.

With the help of JMP's excellent exploratory analysis function, Chicago police have done a lot of analysis work on these data, such as exploring the behavior pattern between the crime and the crime location, and analyzing the crime type based on the time node. Based on the above data analysis, Chicago police carried out a lot of crime prevention work, which greatly reduced the occurrence of crime and the harm brought by crime.

The primary task of data preparation is to sort out the data based on the type of crime. For similar crimes, more detailed grouping will be carried out according to the degree of harm. These are necessary preparations for accurate analysis. Due to the powerful graphical analysis function of JMP, this project not only uses the graphical analysis of JMP in the analysis stage, but also uses the rapid mapping of JMP in the data preparation stage. For example, too much classification of crime types increases the difficulty of identifying the trend. At this time, it needs to be adjusted based on the classification to make the type of crime more reasonable.

By cleaning, deleting and integrating the data, we get all the standardized data needed for analysis, which will greatly help our subsequent analysis work.

JMP's rich graphical analysis tools help Chicago to quickly analyze various criminal incidents. In this project, the police used a lot of graphical analysis work, such as bar chart, line chart, contour map and so on.

For example, the police use bar charts and contour maps for analysis based on crime types and geographical areas. By using the histogram, the police can classify various types of crimes, so that they can clearly see the situation of different types of crimes.

2.3. Case Semantic Region Recognition Based on the Judicial Domain

We propose a new network called Law-net for case semantic region recognition [4]. Law-net is a deep learning network based on full convolution network and semantic segmentation applied to legal texts. In the field of computer vision, full convolution network (FCN) is a well-known

ISSN: 2688-8645

segmentation network. Law-net is a more popular network. Basically, for all segmentation problems, we can take Law-net to look at the basic results first, and then make "magic changes". Because its network structure is roughly U-shaped, it is called Law-net. The network structure is as follows.

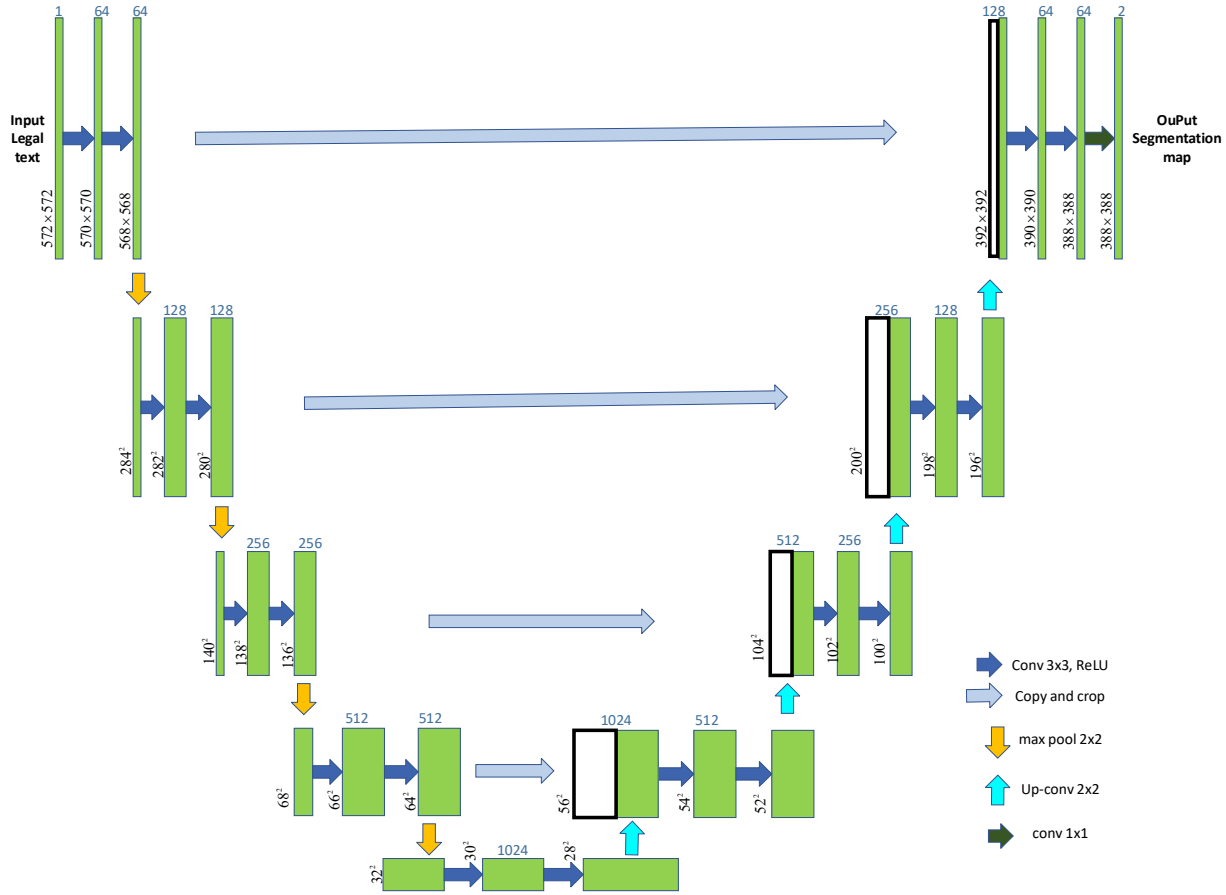

Figure 1. The network structure for Law-net

As shown in Figure 1, full convolution neural network (FCN) is used. The difference between full convolution neural network (FCN) and convolution neural network (CNN) is that FCN replaces the last full connection layer (FC) of CNN with convolution layer (conv), so FCN can input any size of soil photos and output pictures, so this is an end-to-end network.

The contracting path of the network on the left in the figure above: use conv and maxpooling.

The expansive path on the right of the figure above: use upsampling to combine the contracting path on the left with pooling's feature map, and then upsampling layer by layer to 392x392 Heatmap.

Finally, after two conv, we get to the final Heatmap. Then we use 1x1 conv to classify. Here we use two neurons, so we get two heatmaps.

Law-net network supports a small number of training models: in the medical direction, the data we can train is relatively small, while the data / targets that need to be detected are relatively large. So we use data enhancement in Law-net network.

Each pixel can be segmented to obtain higher segmentation accuracy. Pixel level image segmentation is the structural characteristic of Law-net, which is caused by its network structure.

In the process of training, high momentum is used[5]. Here, high momentum = 0.99. The use of high momentum can make the learning rate of training model in a process of self-regulation.

ISSN: 2688-8645

Weight division. In order to make some pixels more characteristic/representative, the weight of pixels is divided to compensate for the different probability of pixels in each image. The formula of weight is as follows.

$$\varphi(\mathrm{x}) = \varphi_c(x) + \varphi_0 \times \exp(-\frac{(D_1(\mathrm{x}) + D_2(\mathrm{x}))^2}{2\sigma^2}). \quad (1)$$

φ_c is a weighting factor used to balance the frequency of categories. D_1 represents the distance to the boundary of the nearest cell, D_2 represents the distance to the boundary of the second nearest cell. Based on experience, we set φ_0 = 20, σ about 5. The weight of our network is initialized by Gaussian distribution, and the standard deviation of the distribution is $\sqrt{2}/N$, where n is the number of incoming nodes.

3. Conclusion

This paper proposes a new method of legal text mining to deal with unstructured legal text, so as to find a lot of knowledge. This method can effectively transform the legal text into a well-structured form, so that big data and artificial intelligence system can generate and verify the model on this basis. However, the function of this method is not complete at present, and the robustness to different legal text formats is not enough. We will improve it in the future work.

Acknowledgments

This work was supported by National Key R&D Program of China (No. 1171-23318102) and Ministry of Education Cooperative Education Project (No. 202002237006).

References

[1] C. Rosca, B. Covrig, and C. Goanta: Return of the AI: An Analysis of Legal Research on Artificial Intelligence using Topic Modeling, Proceedings of the Natural Legal Language Processing Workshop 2020 co-located with the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020), Virtual Workshop, (Online, Aug 24, 2020). Vol. 2645(2020), p.3-10.

[2] P. Boniol, G. Panagopoulos, C. Xypolopoulos: Performance in the Courtroom: Automated Processing and Visualization of Appeal Court Decisions in France, Proceedings of the Natural Legal Language Processing Workshop 2020 co-located with the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020), Virtual Workshop, (Online, Aug 17, 2020). Vol. 2645(2020), p. 11-17.

[3] Li X, Li S, Qin H, et al. Spatiotemporal consistency-based adaptive hand-held video stabilization. Sci China Inf Sci, 2020, 63(1): 114101, https://doi.org/10.1007/s11432-018-9764-0.

[4] Li, T. , et al. "AnaSearch: Extract, Retrieve and Visualize Structured Results from Unstructured Text for Analytical Queries." WSDM '21: The Fourteenth ACM International Conference on Web Search and Data Mining ACM, 2021.

[5] Yates, A. , Nogueira, R. , & Lin, J. . (2021). Pretrained Transformers for Text Ranking: BERT and Beyond. WSDM '21: The Fourteenth ACM International Conference on Web Search and Data Mining. ACM.