

Legal Aid System based on CNN Model

Yufan Zhang ¹, Shugang Liu ²

¹ School of Computer Science, North China Electric Power University, Baoding 071003, China;

² School of Computer Science, North China Electric Power University, Baoding 071003, China.

*Corresponding author Email: 511046268@qq.com

Abstract: China's current rural rule of law construction began after the reform and opening up. Although some achievements have been made, there are still some question [1]. In order to solve the problem of legal inquiry, this paper designs a legal aid system based on CNN model which includes data acquisition module, voice input module, data preprocessing module and semantic similarity calculation module. The results show that the accuracy of the system is 85%, which provides help for people to query legal issues.

Keywords: Deep learning, CNN, legal aid, NLP.

1. INTRODUCTION

With the development of society, people's life is full of information. From communication in everyday life to coding in electronic software, natural language is an important means of expression. People use natural language to describe objects, express themselves and understand each other. From machine translation and artificial intelligence research in 1950s, NLP (Natural Language Processing) has a history of half a century [2]. The application of NLP improves people's ability to process information.

The importance of law in people's life and social development is self-evident. Rural areas are the basic unit of national governance. Rural governance is an important part of the national governance system and a basic project to modernize the national governance system and capacity [3]. In the remote areas, with the lack of law firms and legal aid centers, it is difficult for people to conduct legal inquiries which cost a lot of time and money. How to provide legal aid better and faster has become an urgent problem to be solved.

2. SYSTEMATIC DESIGN

This system mainly includes four functional modules legal which is inquiry, new crown topic, legal recommendation and accusation prediction as is shown in Fig.1

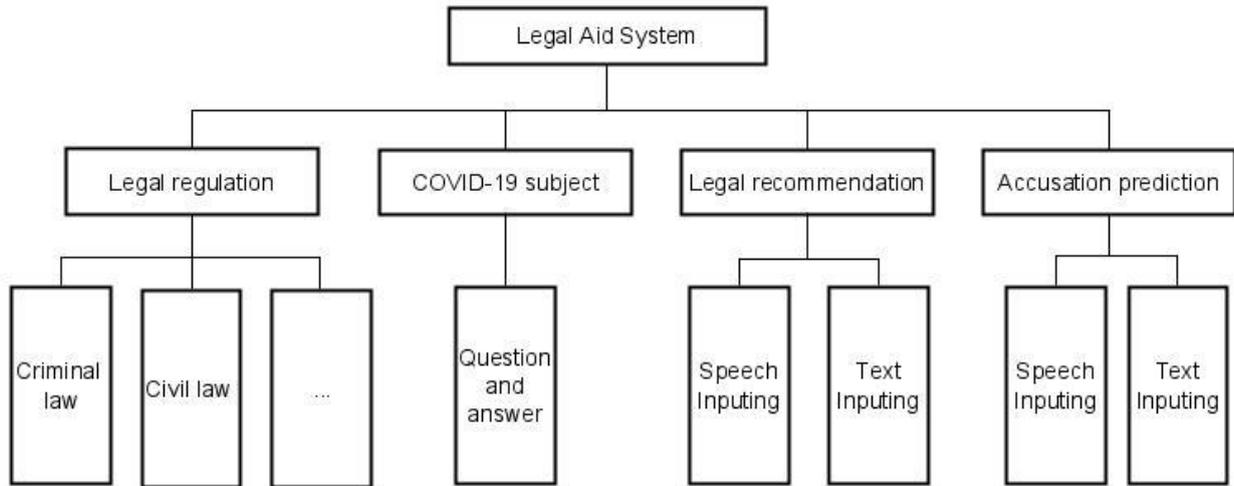


Fig.1 Functional Module Diagram

Among them, the legal regulation mainly does the query of common laws, and displays the specific name and detailed content of each law. In light of the current situation, 23 common questions about COVID-19 have been sorted out from the Office of Legal popularization. The module of legal recommendation and crime prediction realizes the input case by using the autonomous training model. In order to realize the system functions, the system adopts structure as shown in Fig.2. With the help of the Flask, the model is deployed on the cloud server to decouple the model from the application. The system provides speech input and text input to achieve better interactive experience.

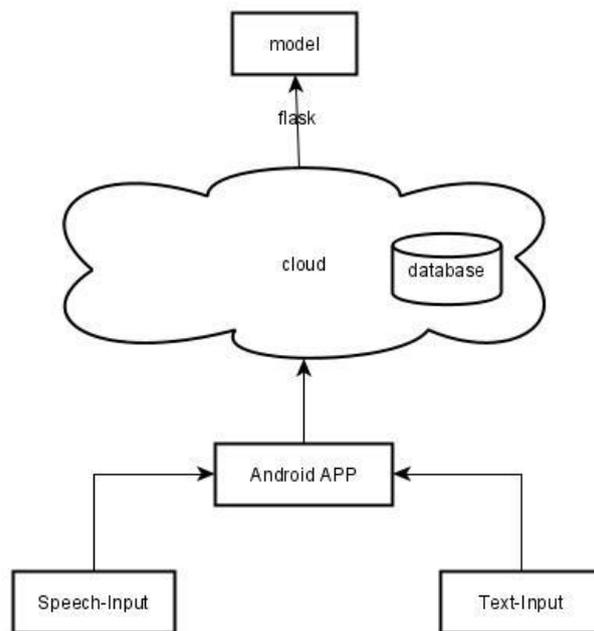


Fig.2 Architecture Design

3. TRAINING PROCESSES

The main processing process includes word vector training, feature vector extraction by the convolutional layer, feature text vector extraction by the pooling layer, and connection feature of the full connection layer, as is shown in Fig.3.

3.1 Word vector training

A case is segmented by using the Jieba, and each word is vectorized by 40,000 dictionary of the word segmentation.

In the training, to remove the model over-fitting, we added the stop word processing. Remove repeated or similar words or expressions that occur in every fact, such as 'someone', 'defendant', 'company', 'branch', 'accuse', 'charges', 'about', 'afternoon', 'morning' and so on.

3.2 The Embedding layer

The Embedding layer converts discrete variables into continuous vectors. While reducing the spatial dimension, these variables are represented more meaningfully

3.3 CNN model convolution layer

CNN model convolution layer is used to extract feature vectors

In the convolution layer, the width of the convolution kernel is n , and the corresponding weight matrix is $h \in R_{n \times m}$. Offset term $b \in R_m$; The activation function used in the experiment is ReLU and denoted by f , then the output of the convolution kernel is (1).

$$T_i = f(h * a + b) \tag{1}$$

Max-pooling is adopted in the pooling layer, as shown in (2).

$$T_i = \max\{T_i\} \tag{2}$$

3.4 The full connection layer

The full connection layer connects the feature vectors. The biggest function of the fully connection layer, the Dense layer, is to transform a feature space linearly to another feature space. After nonlinear changes, the Dense layer to extract the correlation between features and finally map them to the output space. Minimize the influence of features on classification. Fully connections are made using the output vector with an inactivation rate of 0.2.

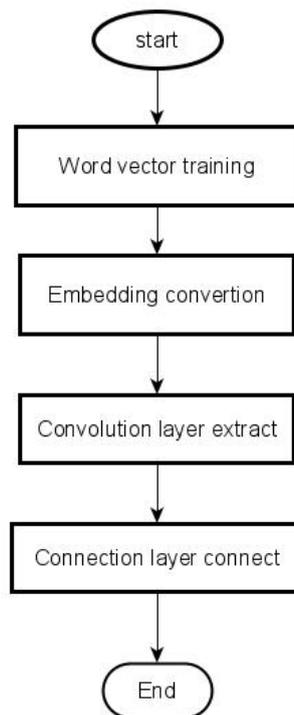


Fig.3 Processing Flow Chart

4. ANALYSIS OF EXPERIMENTAL RESULTS

The experimental results are shown in Table 1 and Table 2. It can be seen that the accuracy rate on the verification set reaches 0.85, indicating that the model can play a certain role in the legal aid system.

Table 1 Test Set Accuracy

Type	Training	Test	Valid
Accuracy	0.9995	0.9982	0.85

Table 2 Valid Set Score

Type	f1_micro	f1_macro	f1_score
Score	0.9072	0.8016	85

According to the training process shown in Fig. 4 and Fig. 5, it can be seen that there is no phenomenon of data fitting when it is trained on the training set, but there is the phenomenon of over-fitting when it is run on the test set. It shows that the generalization ability of the model is not very strong, and it can be adjusted by random training and adding regular term method.

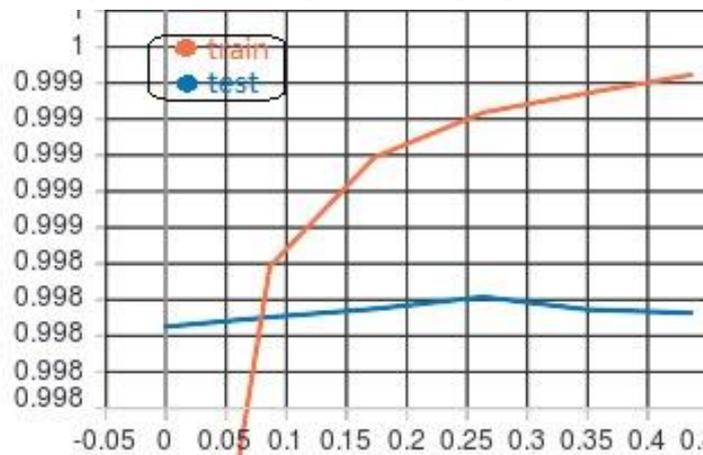


Fig 4. Accuracy of Training Set

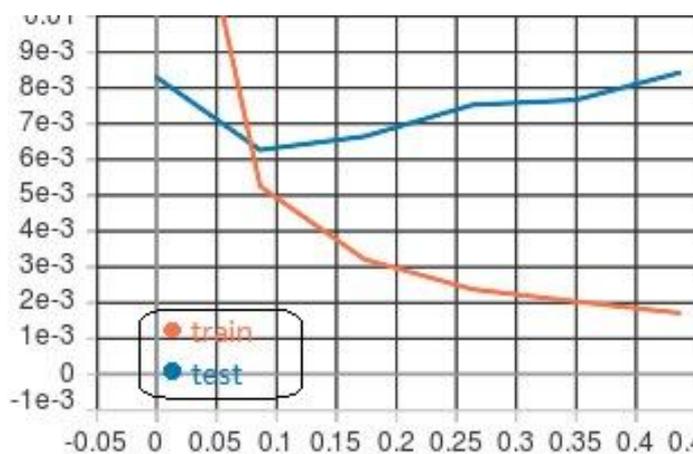


Fig 5. Training Loss Function

5. SUMMARY

The CNN model performs well in the legal recommendation and accusation prediction based on case input. The legal aid system developed based on CNN has a promotion significance for the existing legal self-examination problems.

ACKNOWLEDGMENTS

This work was financially supported by the National Innovation Program-202110079033, Law – A Legal Aid System.

REFERENCES

- [1]. Rui Lian, Penghuai Gao. The Confusion of "Law Not Going to the Countryside" and the Solution -- Also on the Re-definition and Reasonable Application of Folk Law [J]. Journal of Liaocheng University (Social Science Edition),2016(01):108-114.
- [2]. Changning Huang, Xiaofeng Zhang. Natural Language Processing Technology: Three Milestone [J]. Foreign Language Teaching and Research,2002,{4}(03):180-187+239.
- [3]. Wencheng Cai. Grassroots Party Organizations and Modernization of Rural Governance: An Analysis Based on Rural Revitalization Strategy [J]. Theory and Reform,2018,{4}(03):62-71.