

Publication Date: 10 August 2024

Archs Sci. (2024) Volume 74, Issue S1 Pages 32-39, Paper ID 2024s16.
<https://doi.org/10.62227/as/74s16>

The Construction of Modern Administrative Law via Data Mining

Song Zhang^{1,*}

¹ Law Department, Jilin Police College, Changchun 130117, Jilin, China.

Corresponding authors: Song Zhang(e-mail: zhangs9797@163.com).

Abstract Early administrative jurisprudence generally experienced a shift from administrative science to legal science, while modern administrative jurisprudence has shifted from judicial review to administrative process centered. In this paper, we introduce feature engineering technology into the construction model of administrative law based on data mining. The prior knowledge is introduced into the model through the artificially constructed effective features, and the neural network method can automatically extract features such as abstract features from the original input text. We propose a deep feature engineering method to combine the advantages of both in the study of relation extraction tasks. Experiments show that our method provides a powerful legal analysis tool, which helps to innovate the conceptual framework and theoretical system of traditional administrative law and establish a modern administrative law system that can explain the real-world administrative process.

Index Terms data mining, modern administration, law

I. Introduction

It has almost become a consensus that administration is produced before administrative law, but administrative law is not a law about administration, but a law that controls administration [1]. Therefore, early administrative law generally experienced a turn from administrative science to legal science, so that it was able to break out of the cocoon and reborn, and gradually gained a self-consistent status as a legal science [2]. But in the context of the modern administrative state (administrative state), the separation of power structure has undergone a quiet shift [3]. In the face of increasingly complex and highly technical regulatory matters, and in the face of increasing risks of various uncertainties, laws often only provide regulatory frameworks and blueprints, while empowering administrative agencies to fill legal gaps, independently shape and choose public institutions. The power of a vast discretionary space for policy [4], [5]. The administration is no longer merely a "conveyor belt", nor is it any longer to follow the rules and regulations of the legislature. Under the tech-necrotic approach, the courts cannot and do not have the ability to review a large number of highly specialized administrative decisions. Instead, they can only respect the policy judgments and individual decisions of administrative agencies. Moreover, even if an administrative decision is supported by the court, and it does not mean that its decision-making process is impeccable, and it no longer contains legal issues and disputes [6]. In a sense, the administrative process can be called the "lifeline" of modern administrative law, which constitutes the flesh and blood and skeleton of

administrative law.

In fact, modern administrative jurisprudence has shifted the node of control from the "downstream" of the administrative process to the "upstream and midstream", and has taken policy, politics, and law as its own variables; The substantive factors in the administrative process are described, analyzed and judged, and attention is paid to the formation and implementation of public policies [7], [8]. It is against this background that famous contemporary public law scholars in the United States pointed out that the traditional administrative jurisprudence centered on the court lacks a realistic understanding of the substantive objectives, consequences, pathologies and causes of regulatory projects, so legislative and administrative officials should pay more attention, because they are the primary designers of administrative law [9]. British administrative law has also gradually shifted from the red-light theory of "watching the fire from the other side" to the green light theory of "giving one's life to fight the fire" [10]. Japanese scholars have also successively put forward the "administrative process theory", which emphasizes the dynamic grasp of administrative activities or the examination of its actual functions, although there may be differences in their positions, and advocates that all laws that appear in the administrative process should be viewed. The phenomenon is dynamically investigated, the problem is pointed out, and the solution is explored [11].

With the continuous development of Internet technology, digitization has gradually penetrated all walks of life, resulting in massive unstructured texts, from which extracting a large

amount of valuable information can generate considerable economic benefits [12]. To extract structured knowledge from unstructured text and make the complex data high-quality and valuable, the academic community is devoted to the research of information extraction technology. As one of the important sub-tasks, relation extraction information has received extensive attention. The relation extraction task was first proposed at the 7th Message Understanding Conference (MUC-7) held by the U.S. Defense Advanced Research Projects Agency (DARPA) in 1998, and it was determined that the purpose of the task is to extract entity semantic relations, the conference greatly promoted the development of relation extraction and other information extraction tasks [13], [14]. Then in 1999, at the Automatic Content Extraction Evaluation Conference organized by the National Institute of Standards and Technology in the United States, the task of relation extraction was refined and standardized [15], and a standard evaluation system was provided, and then the public information extraction data set ACE was released. , including training corpus and test corpus, officially opened the door to the study of relation extraction tasks. The early research on relation extraction was mainly carried out in the method based on rules and features, which required a high degree of linguistic knowledge of researchers. This kind of method mainly uses artificially defined grammar rules to match the rules in the target text, to obtain the potential entity-relation triples in the text [16]. The advantage of feature engineering is that it can introduce prior knowledge and effectively obtain sentence structure features, but it has poor ability to automatically learn semantic information about target entity pairs in texts. To solve this problem, a deep learning method that can automatically learn abstract features from data and simulate data laws is proposed. As the main implementation technology of deep learning, neural network model can learn and simulate its laws from training data to achieve the purpose of predicting potential features and semantics in new data [17], [18].

In recent years, neural network technology has made great achievements in natural language processing and other fields and has become the main technical means of relation extraction task research. However, the traditional neural network uses the original text as the model input in the research of relation extraction, and the location of the target entity pair in the text and its sentence structure features are insufficiently obtained, which makes the performance of the relation extraction model difficult to improve.

Through the work of this paper, entity relationships have been more accurately identified, providing data support for the development of downstream tasks in the field of natural language processing such as event extraction, knowledge graph construction, and automatic question answering, enriching the theoretical knowledge of natural language processing, and further promoting information extraction technology The development of information can promote the structure and value of complex information. Through the verification of the validity of the feature engineering model, we introduce the feature engineering technology based on data mining into

the construction model of administrative law and realize the instantiation of the validity of the task transfer. Experiments show that our method far outperforms the baseline.

II. Chorus Melody Recognition Algorithm

A. Entity feature construction method

In the traditional relation extraction research based on neural network, the input of the model is the entire sentence, and no special treatment is performed on the target entity pair, which causes the neural network to convert each word in the input text into a corresponding representation indiscriminately, which cannot be obtained. Information about the location and semantics of the target entity. But the target entity in the text has precise location, this chapter can make the neural network focus on the target entity pair by inserting entity features at the boundary of the target entity pair to indicate the entity. When obtaining the text representation, each structurally consistent entity feature is transformed into a vector representation, which is regarded as the structural signature of the relation instance.

After the target entity in the sentence indicates the entity structure on both sides, each entity feature will be mapped to a vector representation of the same dimension as the vocabulary in the sentence. This chapter makes the neural network "attention" to the entity pair in the sentence. location, structure, semantics, and other information. This chapter does a systematic study on the structure of entity features and divides them into three types: entity location features, entity semantic features, and entity comprehensive features. These three types of entity features are briefly described below.

- 1) Entity position feature: This type of feature is used to indicate the position of the target entity pair in the text, making the sentence structure feature centered on the entity pair more obvious. This section divides entity location features into three subcategories: 1) All entity location features are represented by the same character, such as ([P], [P], [P], [P]), where "[P]" is used to indicate entity boundaries;
- 2) Different characters are used for the start and end positions of entities, such as ([P], [/P], [P], [/P]), where "[P]" and "[/P]" marks the start and end positions of the entity respectively; 3) In order to distinguish the two target entities in the relationship instance, use different start and end tags on both sides of the two entities, such as ([P-1], [/P-1], [P-2], [/P-2]), marking the start and end positions of entity 1 and entity 2, respectively.
- 3) Entity semantic features: entity type and subtype features contain important semantic information, so using such features to indicate target entities can capture entity location information and entity semantic information at the same time. This chapter combines entity type with subtype information and entity location features to construct entity semantic features.

This class of features is divided into two subclasses.

Type	Subtype	Formalization	Example
Location features	Dull Positions	P-D	[P], [P]
Location features	Two-side Positions	P-TS	[P], [/P]
Location features	Two-side-ARG Positions	P-TSA	[P-1], [/P-1]
Semantic features	Two-side Types	S-TS-T	[PER], [/PER]
Semantic features	Two-side Subtypes	S-TS-S	[IND], [/IND]
Semantic features	Two-side-ARG Types	S-TSA-T	[PER-1], [/PER-1]
Semantic features	Two-side-ARG Subtypes	S-TSA-S	[IND-1], [/IND-1]
Comprehensive characteristics	Dual- Types-side- ARG	C-DTSA	[PER-IND-1], [/PER-IND-1]
Comprehensive characteristics	Types-POS-side- ARG	C-PTSA	[V-PER-1], [/V-PER-1]

Table 1: Entity characteristics

- 1) Consists only of entity types or subtypes, such as [PER], [/PER], [ORG], [/ORG] or [IND], [/IND], [GOV], [/GOV];
- 2) It consists of entity types or subtypes and features that distinguish the relative positions of two entities, such as [PER_1], [/PER_1], [ORG_2], [/ORG_2]. (Where PER represents the entity type person, ORG represents the entity type organization, IND represents the entity subtype individual, and GOV represents the entity subtype government.)
- 3) Entity comprehensive features: After synthesizing entity location and entity semantic information, this chapter introduces entity-adjacent part-of-speech information to construct entity comprehensive features. This type of feature can encode more entity-related syntactic and semantic information, which is helpful for the neural network to perceive and recognize the target entity pair. In this type of entity feature, we use V to indicate that the part of speech of the adjacent word of the entity is a verb, and N to indicate a noun, etc., which can be expressed as: [V_PER_1], [/N_PER_1], [N_ORG_2], [V_ORG_2]. For the acquisition of part-of-speech information in this part of entity comprehensive features, third-party NLP tools - JIEBA and NLTK are introduced. Finally, this part systematically proposes a total of nine entity features in three categories, as shown in Table 1.

B. BERT-CNN model embedding entity features

After embedding entity features into relation instances, we use vector matrices to obtain text representations, which are then processed by deep neural Network processing, simulating the laws between data. This chapter constructs a neural network model BERT-CNN that combines CNN and BERT models to evaluate the performance of entity features on relation extraction tasks. The model consists of an input layer, an Embedding layer, a convolution layer, a pooling layer, a fully connected layer, and an output layer, and its structure is shown in Figure 1.

Input layer: Four entity features indicating the target entity pair are embedded in the original text to obtain the Positional, semantic and syntactic features. $make = 1, \dots, i, \dots, i + 1, j, \dots, j + 1, \dots, N$ represents a sentence with two named entities, representing the i two words in the sentence, entity 1 = $\dots, i, \dots, i + s$, Entity 2 = $j, \dots, j + t$.

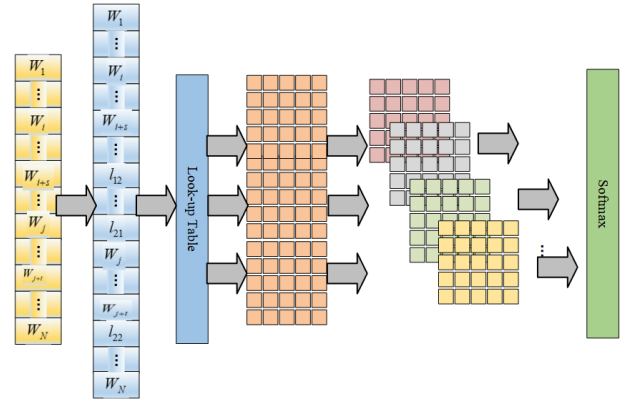


Figure 1: BERT-CNN model

Embedding entity features into S in this layer is used to indicate entity pairs. Due to the symmetry of relation instances, the same entity pair can generate two relation instances I_1 and I_2 , and use $I_1 1, I_1 2, I_2 1, I_2 2$ to indicate the target entity 1 and entity 2 respectively, then we get:

$$\begin{aligned}
 I_1 &= \langle r, e_1, e_2 \rangle \\
 &= w_1, \dots, l_{11}, w_i, \dots, w_{i+s}, l_{12}, \dots \\
 &\quad , l_{21}, w_j, \dots, w_{j+t}, l_{22}, \dots, w_N.
 \end{aligned} \tag{1}$$

$$\begin{aligned}
 I_2 &= \langle r, e_2, e_1 \rangle \\
 &= w_1, \dots, l_{21}, w_i, \dots, w_{i+s}, l_{22}, \dots \\
 &\quad , l_{11}, w_j, \dots, w_{j+t}, l_{12}, \dots, w_N.
 \end{aligned} \tag{2}$$

Embedding layer: three strategies are used to initialize the word vector lookup table, 1) use random numbers to initialize the lookup table; 2) use wiki-100 or GoogleNews-vectors-negative-300 two pre-trained word vector models to initialize Chinese and English word vector lookup table; 3) Use the pre-trained language model BERT to generate a word vector lookup table; convolution layer: In this model, four $3 * 1$ convolution kernels are used to convolve on the output matrix of Embedding, and output 50-dimensional vector. The convolution operation can automatically learn high-level abstract features from the original text; A small amount of efficient feature integration, and finally use the cross-entropy loss function to calculate the loss in the training process; output layer:

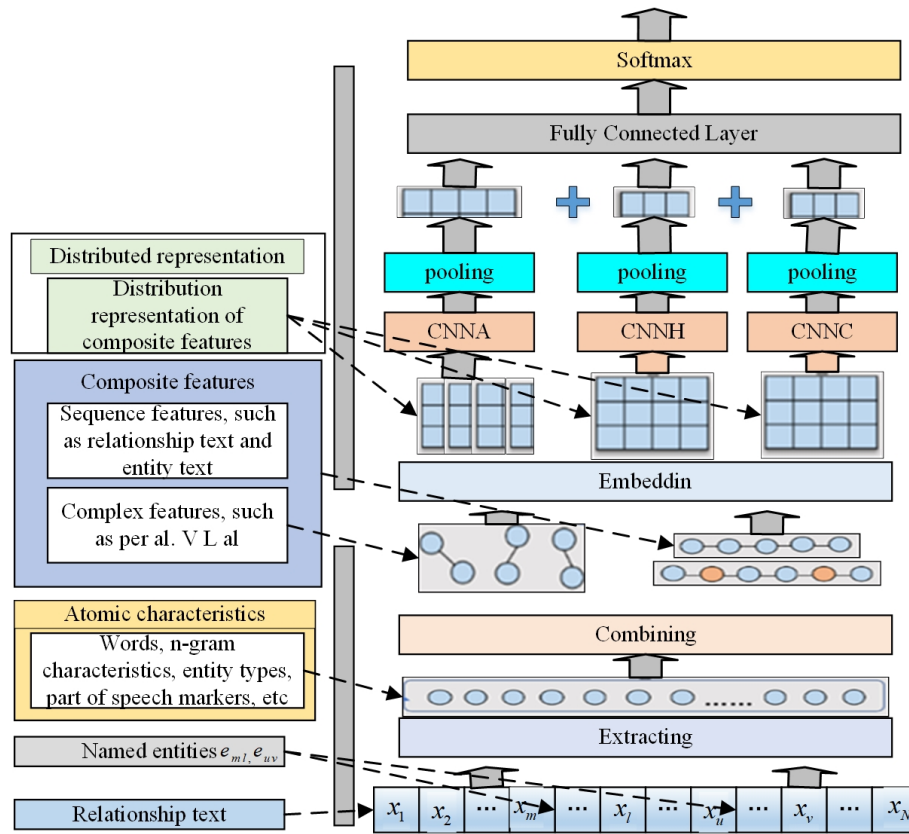


Figure 2: Neural feature fusion model

Convert the output of the neural network through SoftMax and use it in the probability output of each category. In the research of this chapter, using a variety of entity-related information to construct entity features not only enables the neural network to obtain sentence structure features, but also integrates semantic information and syntactic information to further improve the performance of relation extraction.

C. Neutralized Feature Fusion Model

Neutralized feature fusion model combines feature engineering and neural network to extract structural features and semantics of text features, thereby alleviating the feature sparse problem in sentence-level relation extraction. The model is mainly divided into two parts: The overall framework of the model is shown in Figure 2.

The feature engineering structure consists of two steps: feature extraction and feature combination. During feature transformation, convolutional layers transform local features into high-level abstract representations. This section designs three convolutional networks with different architectures (“CNN A”, “CNN H” and “CNN C”) to deal with different types of composite features respectively. Then feature selection is performed through the pooling layer, which captures the features with obvious classification ability from the input abstract features and connects the output of each part into a vector for feature fusion. All inputs are then globally conditioned by a

fully connected layer. Finally, the output is normalized to a probability value using SoftMax to get the classification result.

III. Experiments

A. Dataset Processing

The ACE 2005 dataset was released by ACE (Automatic Content Extraction). The Chinese dataset contains 633 documents. According to the characteristics of the relation extraction task and the experimental requirements of this chapter, the irregular documents 1 were filtered out, and a total of 628 documents were collected. For the relation extraction task, the dataset is manually marked with 9244 relation instances, which are divided into 6 relation categories and 18 relation sub-categories, which are called positive examples. Due to the symmetry of the entity relationship, each entity pair in the sentence should be recognized for the relationship. This chapter extracts sentences containing entity pairs from all documents and regards the instances that do not appear in the positive examples as negative examples. According to this rule Got 98140 negatives. The positive data marked in the ACE 2005 dataset is stored in the apf.xml file, and the negative data needs to be obtained from the sgm file using certain rules and compared with the positive data. The processing process is shown in Figure 3.

First, sentences with two or more entities are extracted from the dataset, and then the entire document storing element

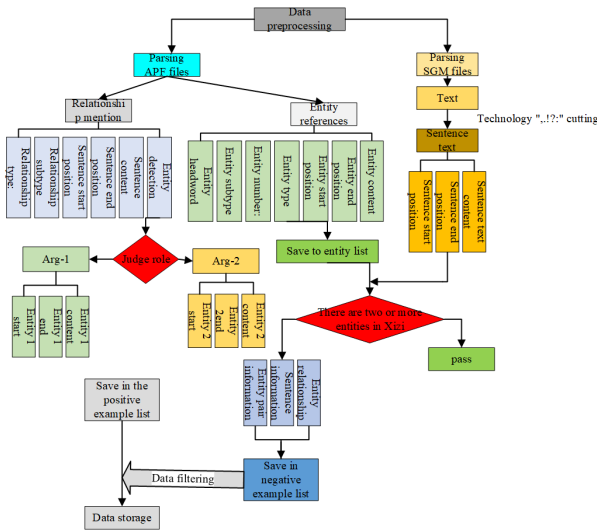


Figure 3: Flow chart of ace2005 data preprocessing

information such as relation instances is subjected to tree analysis, and the sentences marked with the entity relation type in the dataset are extracted as positive examples. use. The interaction with the entire document in data processing is done at the level of tree and element nodes, from which information such as sentence text, entity content, entity type, entity center word, and relation type is extracted. After obtaining the whole article, we use five Chinese (or English) punctuation marks in five Chinese (or English) formats to split the article into sentences and take the instance text that contains entity pairs and does not appear in positive examples as negative examples use. After screening, the experimental data finally obtained in the ACE 2005 Chinese dataset has 107,384 instance texts containing two or more entities. The ACE English dataset and Chinese dataset adopt the same storage and processing methods. There are 506 articles in the English dataset, and finally 6583 positive data and 97534 negative data are obtained from the English dataset.

The positive and negative data are merged to obtain all the instances used for relation extraction in the data set, scramble them into the same file, and divide them according to the needs of the experiment. Three characteristics of the ACE 2005 dataset make the study of relation extraction tasks extremely challenging:

- 1) a large number of negative examples bring serious data imbalance;
- 2) when a negative example and its corresponding positive example only have entity pairs of positions However, the two instances share the same text content, and it is difficult to identify entity relationships through contextual information;
- 3) The dataset contains a large number of nested entities. In the experiment of this chapter, the ACE 2005 dataset is divided into training set, validation set and test set according to the ratio of 8:1:1.

Chinese Literature Text Data Set (CLTC) CLTC is a Chinese

data set released by the MOE Computational Language Key Laboratory of Peking University for two information extraction tasks of named entity recognition and relation extraction. The text of the data set comes from Chinese language literature, it has strong domain language characteristics and aims to solve the problem of lack of information extraction data sets in the field of Chinese literature. The dataset annotates 7 entity types and 9 relation types and annotates relation types under 4 categories of entities: Thing, Person, Location, Organization. Compared with the ACE dataset, the CLTC dataset has no negative samples, no serious data imbalance problem, and the number of nested entities is small, accounting for only about 10%. The dataset file is divided into two types: ".ann" and ".txt", which store the tag information and chapter text of entities and relationships, respectively. The processing steps of this dataset in this experiment are shown in Figure 4.

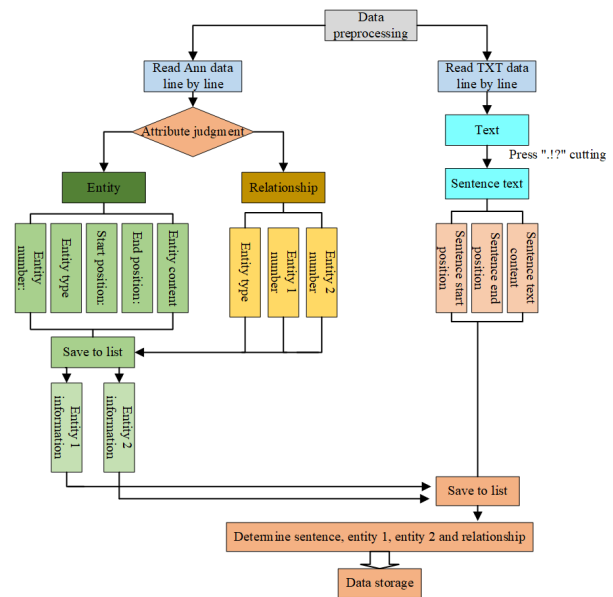


Figure 4: CLTC data preprocessing flow chart

During data preprocessing, the Ann annotation file and the txt source text file of an article are read at the same time and processed accordingly. First read in the txt source text line by line, concatenate sequentially, and get a string that stores the entire document. Divide the chapter into sentences according to the three Chinese punctuation marks of ".!?.", record the starting and ending positions of the sentence in the chapter, and store it in the list S for subsequent detection of the sentence where the entity pair is located. At the same time, the Ann annotation file is read line by line for processing, and the relationship and entity are stored separately for the read data. Store each entity id and entity related information in the list X, which is used to detect the two-entity information corresponding to each labeling relationship, including entity type, entity start and end position, entity content, etc. Match the sentence where the entity pair is located from the list S through the entity start and end positions, that is, get the entity relationship, the target entity pair, and all the information of

the sentence where the entity pair is located, and store it in the file. The final result is 13462 training instances, 1347 validation instances, and 1675 test instances.

B. Experimental parameter settings

This experiment is based on the Python language, the TensorFlow deep learning framework, and is carried out on the Nvidia Tesla P40 platform under the Linux system. Through the analysis of text length in ACE 2005 Chinese dataset and CLTC dataset, the fixed input length parameter is set to 180 and 100, respectively, in the experiment. Experiments show that such parameter settings have a good impact on relation extraction performance. Model training selects cross entropy + L2 regularization as the loss function, Ad delta as the algorithm optimizer, ReLU as the activation function, and Dropout=0.5 is used to prevent overfitting. The remaining parameters are presented in Table 2.

Parameter name	Parameter value
Word vector dimension	769
Learning rate	1.2
Batch quantity	33
Number of iterations	33
Convolution kernel size	3*1
Convolution layer number	5

Table 2: Experimental parameters

Table 2 shows the parameter settings in the BERT-CNN model. When the Embedding layer of the BERT model is not used, wiki-100 is used to map Chinese words, and the word vector dimension is 100; GoogleNews-vectors-negative-300 is used to map English words. The word is mapped, and the word vector dimension is 300.

C. Experimental comparison and result analysis

To comprehensively evaluate the ability of deep relation extraction combined with entity features to obtain sentence structure features, three groups of experiments A, B, and C are set up in this chapter, using the same dataset division scheme and model performance evaluation criteria. Experiment A systematically and comprehensively conducted experimental evaluations on 9 entity features and a group of control methods on three datasets under the same model parameter settings to determine the ability of different groups of entity features to obtain sentence structure features under the deep learning model. , and find the best performing entity feature. In experiment B, the deep entity feature method in this chapter and two existing entity indication methods (position feature PF and multi-channel network model MC-CNN) were compared on three datasets respectively to determine the effect of entity feature method. better. Experiment C-1 applies the best entity features screened in Experiment A to the CNN and BERT-CNN models and compares and analyzes with other Chinese entity relationship extraction models; The best entity feature is applied to the BERT-CNN model, and it is compared and analyzed experimentally with other English entity relation extraction models. (1) Experiment A: Entity Feature Experiment

A conducted experiments on all the entity features listed in Table 1 on three datasets, and their performances are shown in Table 3.

This part of the experiment adopts the CNN network model and initializes the vectors for the Chinese and English datasets with wiki-100 and GoogleNews-vectors-negative-300, respectively. As shown in Table 3, entity features encode more entity-related information, capture sentence structure features and semantic features, thereby improving performance. When using part-of-speech (POS) tags in the "C_PTSA" feature, the experimental performance on the ACE 2005 Chinese dataset is significantly improved, and its F1 value exceeds that of "C_DTSA" by 12%, and the CLTC data Performance on sets has also been improved. On the other hand, in the ACE 2005 English dataset, the performance of using POS tags is degraded, and the reason for this difference may be that English is an alphabetic language, where many adjacent words are function words (e.g., prepositions, pronouns, etc.), which has almost no word sense, causing interference and degrading extraction performance. The results of experiment A show that the "C_PTSA" entity feature, which contains various information such as entity, entity pair structure and part of speech, has achieved the best performance in the Chinese relation extraction experiment, and the "C_DTSA" entity feature containing various information such as entity, entity pair structure and so on has achieved the best performance. "Entity features achieve the best performance in English relation extraction experiments.

On the other hand, because the semantics of Chinese literature text sentences are usually expressed in subtle and special ways, the sentence structure is relatively complex and flexible, and it is more difficult to obtain sentence structure features, so the performance of entity relation extraction on the ACE 2005 Chinese dataset is higher than that of the CLTC data. Set better. This requires that more effective text structure acquisition methods can be proposed according to the language characteristics of this type of text in subsequent research. (2) Experiment B: Comparison with other entity indication methods This part of the experiment uses four methods of CNN, position coding (PF-CNN), multi-channel (MC-CNN) and entity feature for relation extraction to compare different entities The difference between the methods in obtaining sentence structure features and improving the performance of relation extraction is indicated, and the effectiveness of the method in this chapter is verified. The experimental results are shown in Table 4 and 5.

In Table 4, "None" means that no processing is performed on the original text in the input part of the model. In the position encoding model, each word has two distance values relative to two entities, and these two distance values are respectively mapped to a 25-dimensional vector, and the position vector and the word vector are fused to extract richer structural features. However, this method only achieves the purpose of obtaining sentence structure features and does not integrate more entity-related semantic features, which is insufficient to alleviate the problem of feature sparseness

Solid features	ACE (Chinese)/P (%)	ACE (Chinese)/R (%)	ACE (Chinese)/F1(/%)	ACE (English)/P (%)	ACE (English)/R (%)	ACE (English)/ F1(%)	CLTC/P (%)	CLTC/R (%)	CLTC/F1(%)
None	71.28	54.93	60.33	69.55	53.44	60.42	50.65	29.28	37.12
P-C	80.12	59.55	68.32	79.51	61.25	69.19	64.39	55.36	59.53
P-TS	77.43	63.27	69.17	80.81	61.37	69.77	62.46	55.37	59.53
P-TSA	78.66	60.03	68.09	80.14	60.82	69.52	70.14	62.78	66.27
S-TS-T	85.28	65.91	74.35	88.02	68.48	76.48	70.14	62.78	66.25
S-TS-S	83.92	68.90	75.06	84.92	68.56	75.86	×	×	×
S-TSA-T	84.19	71.41	77.32	84.94	70.14	76.83	75.24	72.33	73.75
S-TSA-S	85.92	67.33	75.44	87.04	69.86	77.50	×	×	×
C-DTSA	85.33	70.71	77.34	86.77	71.38	78.32	×	×	×
C-PTSA	91.19	88.56	89.86	85.93	69.92	77.12	75.92	73.58	74.73

Table 3: Experimental performance of entity characteristics

Data set	None/P (%)	None/R (%)	None/F1(/%)	PF-CNN/P (%)	PF-CNN/R (%)	PF-CNN/ F1(%)
ACE (Chinese)	71.31	54.93	60.33	65.02	59.24	61.98
ACE (English)	69.55	53.44	60.41	77.82	59.71	67.54
CLTC	50.65	29.28	37.12	54.77	37.33	44.38

Table 4: Performance comparison of entity indication method 1

Data set	MC-CNN/P (%)	MC-CNN/R (%)	MC-CNN/F1(/%)	Solid features/P (%)	Solid features/R (%)	Solid features/ F1(%)
ACE (Chinese)	73.57	53.94	61.98	91.19	88.56	89.87
ACE (English)	80.29	60.76	67.77	86.78	71.38	78.32
CLTC	69.25	64.38	66.75	75.92	73.58	74.73

Table 5: Performance comparison of entity indication method 2

Serial number	Model	Features	F1(%)
1	SVM	Word vector, entity, WordNet, How Net, part of speech	48.8
2	RNN	Word vector, part of speech, entity, WordNet	49.2
3	CNN	Word vector, position vector, entity, WordNet	52.5
4	CR-CNN	Word vector and position vector	54.2
5	SDP-LSTM	Word vector, part of speech, entity, WordNet	55.4
6	Dep NN	Word vector, WordNet	55.3
7	BRCNN	Word vector, part of speech, entity, WordNet	55.7
8	C-ATT-BLSTM	Character vector, position vector and entity information	56.3
9	SR-BRCNN	Word vector, part of speech, entity, WordNet	66.1
10	Random-CNN	"S - TSA - T" solid feature	74.73
11	BERT-CNN	"S - TSA - T" solid feature	77.15

Table 6: Model performance comparison (CLTC)

and target entity focus confusion. In the multi-channel neural network model, each relation instance is divided into five parts by entity pairs, and then each part uses an independent lookup table, The representation of the same word in a relation instance but in different channels can be obtained. Since the entity pair contains less contextual semantics in the channel, it can play a role in highlighting the position and structure of the entity pair, but it still does not integrate more semantic features.

The ability to perceive the target entity pair is weak; through experiment A, it is known that the "C_DTSA" feature has the highest extraction performance on the ACE English dataset, and the "C_PTSA" feature has the highest performance on the ACE 2005 Chinese dataset and CLTC dataset. performance, so it is selected correspondingly in experiment B. In the entity feature method, this experiment inserts the "C_PTSA" feature including entity type, part of speech, entity pair feature into the instance text of the Chinese dataset (ACE 2005 Chinese and CLTC) to indicate the target entity pair, and the "C_DTSA" Features are used to indicate target entity pairs in instance texts in the ACE 2005 English dataset. The experimental results in Table 3 and Table 5 show that the deep relation extraction method combined with entity features in this chapter can significantly improve the extraction per-

formance on both Chinese and English datasets. Positional encoding PF-CNN and multi-channel neural network MC-CNN are two deep neural network models used to obtain sentence structure features in relation extraction. By analyzing and comparing the experimental data obtained in Table 3, 4 and 5, these two methods can greatly improve the performance of entity relation extraction on the CLTC dataset, but the improvement effect is not obvious enough on the ACE 2005 dataset. This is because the CLTC dataset is different from the ACE 2005 dataset, which contains many negative samples and has a serious data imbalance problem. And since there are many entity-to-symmetric position relation instances in the ACE 2005 dataset, each positive relation instance corresponds to a negative instance with the same contextual semantics. Therefore, the location information of named entities is more complex, resulting in the inability to greatly improve the performance. Comparing these three methods with the deep relationship extraction method combined with entity features, it can be seen from the experimental results that the deep relationship extraction method combined with entity features can achieve very good results in multi-language and multi-domain, this is because Combining the entity feature method can enable the neural network to encode the target entity location information, sentence structure information, text se-

mantic information, syntactic information and other important features for the study of entity relationship extraction, and alleviate the problem of target entity focus confusion.

This part of the experiment is also verified on the CLTC Chinese relation extraction dataset. The publisher of this dataset has performed comparative experiments on many popular relation extraction models on the dataset. The performance is shown in Table 6. It can be seen that our methods can obtain a good performance.

IV. Conclusion

In this paper, the vocabulary and language in the administrative law text are extracted by data mining, and the relationship extraction text contains multiple entity pairs. The traditional neural network-based relationship extraction uses the original text as the model input and cannot obtain the location and semantics of the target entity pair. We propose a deep relationship extraction method combining entity features to construct entity features and placed on both sides of the target entity in the text to obtain sentence structure features and text semantic features. The experimental results show that the deep relation extraction method combined with entity features can stably improve the relation extraction performance under CNN, BERT, or BERT-CNN models. In view of the validity of the experiment in many texts, this provides an important basis for the follow-up research on the construction of modern administrative law.

Funding

There is no funding support for this study.

References

- [1] Dai, H. , Hu, X. , Shan, Z. , & Liu, X. . (2021). Research on the guidance relationship construction of modern apprenticeship in enterprises based on seci model. *Journal of Physics Conference Series*, 1744(3), 032097.
- [2] Ding, X. , & Hu, H. . (2021). On the campus construction environment under the background of modern quality education. *Open Access Library Journal*, 8(9), 6.
- [3] Wang, X. . (2021). Construction of civil engineering teaching system based on data mining algorithm and big data technology. *Journal of Physics: Conference Series*, 1852(3), 032022 (6pp).
- [4] Li, Y. . (2021). Research on the construction of college english mixed teaching model based on modern educational technology and computer technology. *Journal of Physics Conference Series*, 1915(2), 022091.
- [5] Zhou, H., Wu, T., Sun, K., & Zhang, C. (2022). Towards high accuracy pedestrian detection on edge gpus. *Sensors*, 22(16), 5980.
- [6] Bahirov, S. . (2021). "inconsequent carelessness" in modern conditions of criminal legislative reform: the problem and ways to solve it. *NaUKMA Research Papers Law*, 7, 3-13.
- [7] Redkov, N. A. , Ganzha, N. V. , & Khlestakova, L. A. . (2021). On the need to improve criminal law norms with administrative prejudice. *Vestnik of Kostroma State University*, 27(2), 175-180.
- [8] Deng, J. , & Qing, H. . (2021). Cross-border e-commerce course construction based on data mining algorithm. *Journal of Physics: Conference Series*, 1852(2), 022040 (7pp).
- [9] Agenbag, H. , & Amoah, C. . (2021). The impact of modern construction technology on the workforce in the construction industry. *IOP Conference Series: Earth and Environmental Science*, 654(1), 012001 (10pp).
- [10] Budyakov, A. S. , Cherckesova, L. V. , Titov, A. E. , & Yu, A. S. . (2021). Prospects for design and construction development of integral circuits of lc-filters based on the modern dielectric materials. *IOP Conference Series: Materials Science and Engineering*, 1029(1), 012098 (9pp).
- [11] Zheng, C. , & Zhou, W. . (2021). Research on information construction and management of education management based on data mining. *Journal of Physics: Conference Series*, 1881(4), 042073 (6pp).
- [12] Grebenyuk, E. A. . (2021). Identification of significant factors for improving the efficiency of enterprises based on the construction of an efficient frontier. *Journal of Physics: Conference Series*, 1864(1), 012101 (8pp).
- [13] Pei, Y. . (2021). Research on the construction of quality monitoring system party building in private colleges and universities based on big data technology analysis. *Journal of Physics Conference Series*, 1744(4), 042103.
- [14] Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M. F. (2020). Government by algorithm: Artificial intelligence in federal administrative agencies. *NYU School of Law, Public Law Research Paper*, (20-54).
- [15] Shu, X., & Ye, Y. (2023). Knowledge Discovery: Methods from data mining and machine learning. *Social Science Research*, 110, 102817.
- [16] Hamdi, A., Shaban, K., Erradi, A., Mohamed, A., Rumi, S. K., & Salim, F. D. (2022). Spatiotemporal data mining: a survey on challenges and open problems. *Artificial Intelligence Review*, 1-48.
- [17] Kushwaha, A. K., Kar, A. K., & Dwivedi, Y. K. (2021). Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights*, 1(2), 100017.
- [18] Shang, H., Lu, D., & Zhou, Q. (2021). Early warning of enterprise finance risk of big data mining in internet of things based on fuzzy association rules. *Neural Computing and Applications*, 33(9), 3901-3909.

...