



Research and Design on Cognitive Computing Framework for Predicting Judicial Decisions

Jiajing Li^{1,2} · Guoying Zhang¹ · Longxue Yu² · Tao Meng²

Received: 17 October 2018 / Revised: 10 November 2018 / Accepted: 28 November 2018 / Published online: 10 December 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

This paper aims to provide a cognitive computing framework to meet the challenges of semantic understanding, knowledge learning and judicial reasoning in the Chinese legal domain. In our framework, legal factors are first represented in a formal way; secondly, legal factors are extracted, and concepts and their relations are augmented with a combination of rule-based and deep learning methods; thirdly, a predication model is generated and trained to make judicial decisions. When a fact description is brought into the model, the probability of judicial decisions will be given automatically. Two elementary results are obtained: I. Our method can effectively predict the decisions for divorce cases with different expression styles, and offers better performance than traditional methods like Support Vector Machine (SVM); II. Our machine learning predicting results can be easily understood by general public as applied induction rules are given.

Keywords Artificial intelligence · Cognitive computing · Natural language processing · Machine learning · Legal science · Judicial decisions

1 Introduction

The task of predicting judicial decisions is to analyze law articles and documents, extract legal factors as well as their relationship from a large number of historical cases, and determine the judicial outcomes for a pending case by analyzing its textual fact description. This work is helpful for general public to foresee the possible outcomes of their interested law cases. Due to the insufficiency of

law knowledge, it is hard for laypeople to understand professional legal terms and judge the law cases, without the help of legal experts.

In recent years, researchers have taken advantage of Natural Language Processing (NLP) and Machine Learning (ML) techniques to a variety of applications in the legal domain. However, predicting decisions of a case based on fact description is not trivial. There are some major challenges for this work as follows:

- (1) Legal words don't have a global standard in Chinese. On one hand, semantic ambiguity is widespread; on the other hand, general public tend to use spoken words to describe law facts, which are different from the legal terms used by legal professionals. For instance, one may use "beating", "injuring", "maltreatment" and so on, to describe the concept of "domestic violence" in divorce cases. Although the pieces of text are different, they actually have the same meaning. This requires computer systems to capture semantics with legal basis.
- (2) There are a large number of out-of-vocabulary words in judicial documents and fact description, especially in civil cases. The traditional Chinese segmentation algorithm cannot recognize them well. So it is crucial for prediction method to learn those new concepts and

✉ Jiajing Li
lijiajing@aliyun.com

Guoying Zhang
zhangguoying1101@163.com

Longxue Yu
yulongxue@tmlsystem.com

Tao Meng
mengtao@tmlsystem.com

¹ School of Mechanical Electronic & Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China

² Nanjing Wangganzhicha Information Tech. Inc., Nanjing 210094, China

relations automatically or semi-automatically, to make legal knowledge augmentation, and gain a better recall.

- (3) The prediction results should be understandable by general public. Machine learning techniques have been successful in a variety of applications in the legal domain, like retrieval and classification of legal documents, acquiring relevant law articles, and automatic charge prediction. These work employed algorithms like Conditional Random field (CRF) [1], SVM [2, 3] and neural networks [4]. But the key problem here is, they failed to answer why the prediction results are correct, and the prediction results are hard to interpret.

In this paper we proposed a cognitive computing system framework for predicting judicial decisions to meet these challenges. In general words, cognitive computing refers to the technology platforms based on scientific disciplines of artificial intelligence and signal processing, which mimic the function of human brain, and help to improve human decision-making [5]. We proposed a three-layer structure for legal semantic understanding, legal knowledge learning, and judicial reasoning. First, legal factors are represented in a formal way; Secondly, legal factors are extracted, and concepts and relations are augmented with a combination of rules and deep learning methods; Thirdly a prediction model is generated and trained. When a fact description is brought into the framework, the probability of each result will be given automatically. Our approach has the following advantages:

- (1) It is based on legal knowledge representation and extraction. The semantic alignment in this background is flexible for different expression styles in fact description. So general public can describe the cases, or express the queries, with their daily or spoken vocabulary.
- (2) It takes advantage of both artificial rule-based method and deep learning in a complementary way. Artificial rules are written to define the judgement logic, as well as the related concepts and relations. So our work doesn't need a large amount of semantic labelling. At the same time, deep learning algorithms such as Bi-LSTM and CNN are used to proceed legal knowledge augmentation and achieve a higher recall.
- (3) Its predicting results are interpretable in a way that induction rules are supplied. This helps laypeople or non-professionals, to understand the judicial decisions.

We evaluate our framework in the context of divorce cases in Chinese. The experimental results show that our method can effectively predict the decisions for divorce cases in different expression styles, and offers better performance than SVM; and the prediction results are interpretable as applied induction rules used by the system are given.

2 Related Work

Cognitive computing incorporates a wide range of approaches in the fields of information analysis, NLP and ML, and helps policymakers uncover extraordinary insights from vast amount of unstructured data [6]. For instance, IBM's Watson has been successfully used in various fields such as finance and healthcare [7]. We supplied a cognitive computing framework to meet challenges of semantics understanding, knowledge learning and reasoning in predicting judicial decisions.

NLP and ML have been provided to a variety of applications in the legal domain in recent years. One related work is the query-based retrieval of relevant legal judgments. Chen et al. [8] introduce a text-mining-based method to help the general public to acquire relevant criminal judgments using ordinary terminology or statements as queries. Raghav et al. [9] proposed an approach to find similar judgments by exploiting citations in legal judgments through clustering. In their follow-up work, an effort has been made to improve the search ranking of returned judgements based on enhancement of similarity between judgement document paragraphs [10].

There are preceding researches that aim at developing systems to answer legal questions. Kim et al. [11] used ranking SVM and syntactic/semantic similarity to extract relevant Japanese Civil Code articles first, and then use them to answer the yes/no questions. Carvalho et al. [12] proposed a method combining lexical and morphological characteristics, to find relevant documents to a legal question, and extract textual entailment evidence to provide a correct answer.

There are also research works focusing on acquiring relevant law articles or status for a given case in the civil law system. Liu et al. [13] employed techniques of instance-based classification and introspective learning for classifying charges of larceny and gambling. Liu et al. [2] proposed an innovative method named TPP (Three Phase Prediction), in which SVM was first used for preliminary article classification, and then word level features and co-occurrence tendency among articles were used to re-rank the results.

The charge prediction task is to determine the correct charges based on the case fact. Liu et al. [14, 15] used KNN to classify criminal charges in Taiwan by using word-level and phrase-level features. Lin et al. [1] manually designed legal factor labels of robbery and intimidation cases for case classification and sentencing prediction; Luo et al. [4] proposed an attention-based neural network method to jointly model the charge prediction task and the relevant article extraction task in a unified framework.

Another related work is to predict the overall decisions of a case. A more recent research shows how to improve

models for predicting the votes of the US Supreme Court judges [16]. Aletras et al. [3] used textual content such as N-grams and topic models to predict which party will be the winning side. Katz et al. [17] used a randomized tree method to predict whether the court will affirm or reverse the decision of a lower court. In our method, we used semantic factors instead of word sequences for the prediction.

3 Cognitive Computing Framework

Cognitive computing focuses on reasoning and learning, and integrating this capability with specific domain knowledge, to solve business problems. In our case, we employ cognitive computing for predicting judicial decisions. As depicted in Fig. 1, our cognitive computing framework has three layers:

- (1) Legal semantics understanding layer

The legal semantic understanding layer aims to represent the semantics of legal factors in judicial

- documents and fact description in a formal way, so as to carry out subsequent learning and reasoning.
- (2) Legal knowledge learning layer

The legal knowledge learning layer extracts legal factors from judicial documents or fact descriptions; furthermore augments concepts and relationships in a combination way of rule-based method and deep learning.
- (3) Legal knowledge reasoning layer

The legal knowledge reasoning layer tries to learn the implicit casual relations between legal factors based on massive judicial documents, and furthermore predict the judicial decisions according to the fact description.

The following sections will describe the workflow.

3.1 Legal Semantics Understanding Layer

The main work includes data preparation, the establishment of first-order logic base, and the extraction rules.

Figure 1 Cognitive computing framework for predicting judicial decisions.

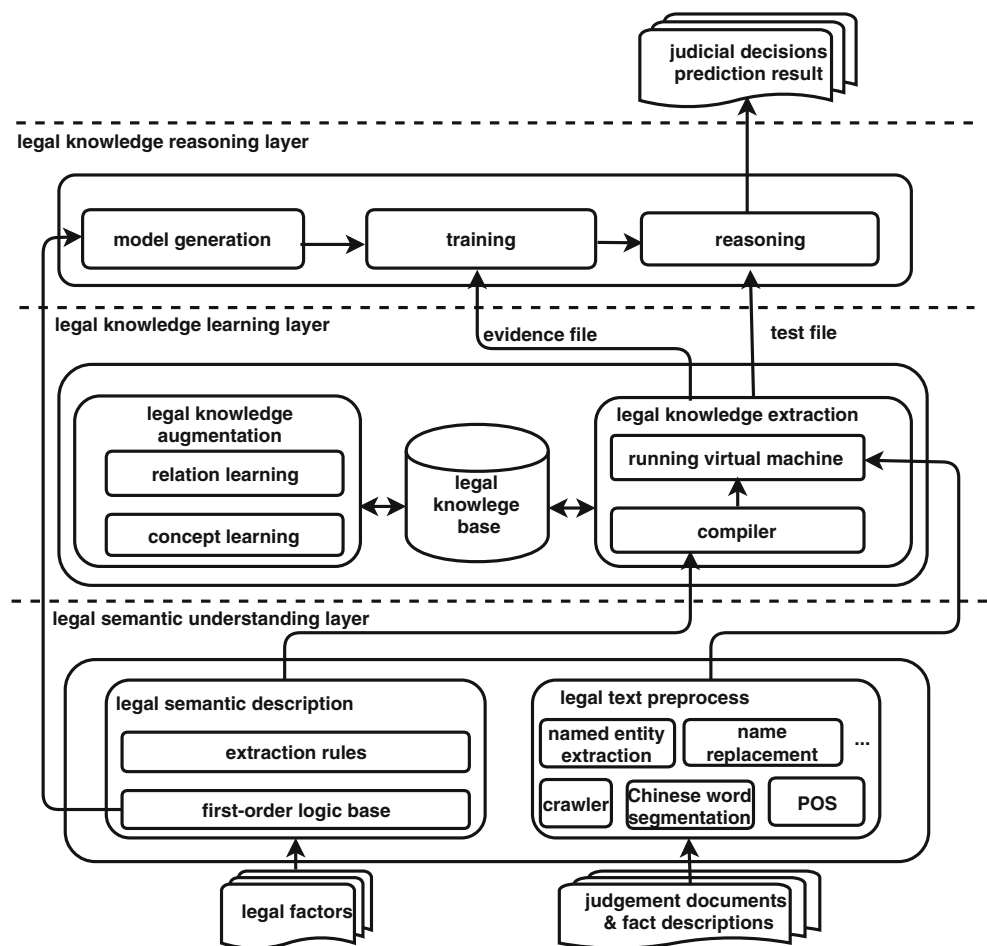
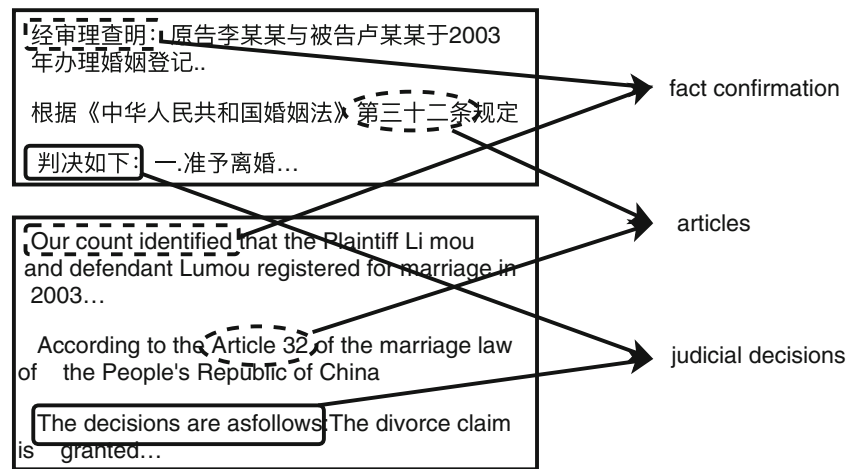


Figure 2 An example of judicial document.



3.1.1 Data Preparation

Our data are collected from China Judgments Online (<http://wenshu.court.gov.cn>). An example of judicial document is shown in Fig. 2, where we highlight the indicator clauses, which we used to divide a document into three pieces, and thus extract fact confirmation, articles and judicial decision, respectively. Fact factors are extracted from the “fact confirmation” piece and result factors are extracted from the “judicial decisions” piece respectively.

The input of judicial documents and fact descriptions are preprocessed by common technologies including Chinese word segmentation, lexical labeling, named entity extraction, text classification, keyword extraction, name replacement and so on. It works with Hadoop to handle massive data.

3.1.2 First Order Logic Base

A first-order logic base is a collection of first order logical statements or rules. There are three types of elements in the first-order logic base: variables, predicates and formulas.

(1) Variables

Variables are summarized and defined according to the legal factors by legal experts. There are two

types of legal factors in our work, namely fact factors and result factors, while the fact factors are the key elements in the fact description, and the result factors are the key elements in the trial result. Legal factors are taken as variables and assigned to different values in various cases.

Taking divorce cases as an example, fact factors include: mutual affection status, the effectiveness of adjust, the number of children and so on. Table 1 shows several examples of fact factor variables.

Result factors include whether the divorce claim is granted, ownership of child custody and so on. Table 2 shows several examples of result factor variables.

(2) Predicate

In our approach, a predicate represents the attribute of legal factor variables. Table 3 shows several examples of predicates.

(3) Formula

The formulas refer to the dependency between fact factors and result factors. Table 4 shows some examples of formulas.

Each formula has a corresponding weight. They are initially set to 0.1. The plus sign “+” means that different values of variables are assigned different weight values.

Table 1 Several examples of fact factor variables.

Fact factor variable	Meaning
$lovestatustype = \{0, 1, 2\}$	“0” means mutual affection broken; “1” means one party still has feelings for the other; “2” means both parties still have feelings for each other
$adjusteffective = \{0, 1\}$	“0” means that out of court adjust is ineffective and “1” means adjust is effective
$childnum = \{0, 1, 2, 3\}$	“0” means childless, “1” means having one child, “2” means having two children, and “3” means having more than two children
$havingbadhabit = \{0, 1\}$	“0” means nobody has a bad habit, “1” means someone has a bad habit

Table 2 Several examples of result factor variables.

Result factor variable	Meaning
$judgmenttype = \{0, 1, 2\}$	“0” means divorce claim is granted, “1” means divorce claim is rejected, and “2” means out-of-court settlement
$custodytype = \{0, 1, 2, 3\}$	“0” means childless, “1” means right of custody belongs to plaintiff, “2” means right of custody belongs to the defendant, and “3” means both parties bring up their children together
$maintenancepayer = \{0, 1, 2, 3\}$	“0” means childless, “1” means the maintenance fee should be paid by the plaintiff, “2” means the maintenance fee should be paid by the defendant, and “3” means the maintenance fee should be paid by both parties
$feepayer = \{0, 1\}$	“0” means the trial fee should be paid by the plaintiff, “1” means the trial fee should be paid by the defendant

Table 3 Several examples of predicates.

Predicate	Meaning
$ Lovestatus(case, lovestatustype)$	the value of <i>lovestatustype</i> in <i>case</i>
$ Haschildren(case, childrennum)$	the value of <i>childrennum</i> in <i>case</i>
$ Isadjusteffective(case, adjusteffective)$	the value of <i>adjusteffective</i> in <i>case</i>
$ Ishavingbadhabit(case, havingbadhabit)$	the value of <i>havingbadhabit</i> in <i>case</i>
$ Judgement(case, judgementtype)$	the value of <i>judgementtype</i> in <i>case</i>
$ Custody(case, custodytype)$	the value of <i>custodytype</i> in <i>case</i>

Table 4 Several examples of formulas.

Formula	Meaning
$ Ishavingbadhabit(x, +y) \Rightarrow Judgment(x, +z)$	the relation between <i>havingbadhabit</i> and <i>judgmenttype</i>
$ Lovestatus(x, +y) \wedge Haschild(x, +z) \Rightarrow Custody(x, +a)$	the relation between <i>lovestatustype</i> , <i>childnum</i> and <i>custodytype</i>

Table 5 Context operators.

Context Computing Operator	Meaning
AND	all clauses in the scope must appear simultaneously in the input text
OR	at least one of all clauses in the scope should appear in the input text
NOT	clauses in the scope cannot appear, otherwise the input text does not match
SENT	all clauses in the scope must appear in a single statement
$Dist_n$	the distance between any two adjacent clauses in the scope cannot be greater than <i>n</i>
ORD	all clauses in the scope must appear in order
CONT	all clauses in the scope must be adjacent

3.1.3 Extraction Rules

Legal factors consist of concepts and their relations. We represent the semantics of legal factors by extraction rules which are written in a programming language named TML [18]. TML is the foundation of our platform which supplies methods like generative grammar and context operators to define semantics and knowledge. In TML, the concepts and relationships are represented in the form of non-terminators, while strings, regular expressions and operators are in the form of terminators. Table 5 specifies our context operators.

An example of concept is shown as following:

CONCEPT BADHABIT:=OR(“gambling”, “drinking”, “taking drugs”,...”).

This extraction rule specifies the concept “BADHABIT” will be matched where at least one of gambling, drinking or taking drugs appears in text.

An example of the relation of concepts is shown as following:

1. CONCEPT ACCUSER-DEFENDANT:=OR(“plaintiff”, “defendant”);
2. CONCEPT FREQUENCY:=OR(“often”, “many times”);
3. RELATION HAVING-BADHABIT(ACCUSER-DEFENDANT who, FREQUENCY fq, BADHABIT bh){
 - ORD(Dist_5(who, fq),bh);

On line 1, the concept “ACCUSER-DEFENDANT” is defined; On line 2, the concept “FREQUENCY” is defined; One line 3, the relation “HAS-BADHABIT” is defined, Where “ACCUSER-DEFENDANT”, “FREQUENCY” and “BADHABIT” should appear in sequence, and the distance between “ACCUSER-DEFENDANT” and “FREQUENCY” should not more than 5 words. Our work of TML is introduced in [18].

3.2 Legal Knowledge Learning Layer

In this layer, the values of legal factors are extracted, and concepts and relations are augmented through deep learning.

3.2.1 Legal Knowledge Extraction

We implemented a compiler and a running virtual machine for TML. In the compiler, extraction rules are compiled and optimized, and machine learning model is trained based on the rule-matching results as labeling corpus. Now given a new piece of text to the running virtual machine, the matched rules will be output as well as the related concepts and relations.

In the semantic pattern matching process, extraction rules which do not contain operators can be directly combined

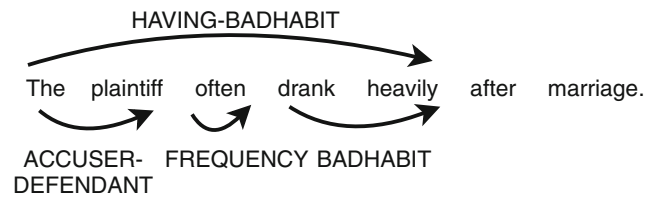


Figure 3 An example of text extraction result.

and transformed into finite state automata for high speed; while rules which contain context operators are converted into a set of running virtual machine instructions and operands. In Fig. 3 we give an example of text extraction result. The details of text extraction are specified in [18].

After extraction, legal factors are instantiated and assigned to concrete values. For example, the text “The plaintiff often drank heavily after marriage” matches the pattern “HAVING-BADHABIT”, so the fact factor variable *havingbadhabit* is assigned to “1”.

3.2.2 Legal Knowledge Augmentation

The result of knowledge extraction from judicial documents can be further used for knowledge augmentation.

(1) Concept augmentation

There are two ways to learn new concepts in our work, one way is to learn from the internal composition of a concept, and the other is to learn from its external context.

The former method takes advantage of the semantic similarity of concepts. For example, each explicitly defined synonym can be found based on word vectors; and concepts which are labeled as synonymous with multiple instances of one concept can be identified as new concepts of the same type. We trained Google’s word2vec model for this step [19].

The latter method uses contextual features. For the output concepts extracted by rules, the matching results are saved in a BMES format. Even without further manual confirmation, they can be used as a sequence role tagging corpus as they are produced from expert written rules, so sequence tagging algorithms like Bi-LSTM and CRF [20, 21] can be used. Bi-LSTM structure provides complete past and future context information for each point in the input sequence of the output layer. CRF makes predictions at the sentence level, so that the probability of the final sequence annotation results is maximized. In these ways we get new concepts.

(2) Relation augmentation

Relation augmentation is to learn the constraints between concepts. The sentences that match the

Table 6 Result of divorce prediction.

Divorce	Probability(%)
Granted	46.55
Rejected	53.45

relation definition are taken as positive examples, and those that don't match the relation are taken as counter examples; then the relation augmentation task can be regarded as a classification problem. We use both Naive Bayes and convolution network (CNN) [22] to accomplish this work.

3.3 Legal Knowledge Reasoning Layer

Based on the first logic base, a Markov Logic Network (MLN) model will be generated. A MLN consists of a series of $\langle F_i, w_i \rangle$ pairs, where F_i is a first order formula and w_i is the weight of it. Then judicial documents are used to train the network and learn the weights of rules. The method we computed a MLN can be found in [23].

4 Experiments

There are 695,418 judicial documents of divorce cases in our system. We randomly chose 50,000 judicial documents, 80% of which were used as training cases and 20% as test cases. Judicial documents and fact description are preprocessed firstly. After that fact factors and result factors that are extracted from judicial documents make up the evidence file, and fact factors are extracted from fact description make up the test file. With the test file used as input and result factors used as query predicates, reasoning is carried out on the trained MLN. The prediction results include text description along with the probability value of each result factor. The operating environment of the experimental system is inter-xeon 2.2 GHz CPU, 8G RAM and Linux operating system.

Here is an example of the fact description:

“The plaintiff and the defendant met and fell in love in 1999. They registered for marriage in October 2003. Their son was born in 2004. After marriage their relation is so so, and they often quarrel about family matters. In 2009, the

Table 8 Result of maintenance payer prediction.

Maintenance Payer	Probability(%)
the plaintiff	39.28
the defendant	60.72

defendant had an extramarital affair, and his attitude to the plaintiff was getting worse and worse. He had beaten the plaintiff several times, and paid no attention to their son.”

Here is the text description of the prediction result:

1. There are contradictions between the couple;
2. There is an extramarital affair;
3. There is a domestic violence.

Further, the predicting judicial decisions are:

1. The divorce claim is rejected;
2. The right of maintenance will belong to the plaintiff;
3. The defendant should bear the maintenance costs;
4. The plaintiff should bear the case fees.”

For the same case, Tables 6, 7, 8 and 9 respectively gives the prediction result of divorce, ownership of custody and maintenance payer. It should be noted that in our work the result factors were considered to be independent of each other, and will be optimized in the future.

We also implemented an SVM model, which is effective and scales well in many fact-description-related tasks in the field of law-AI [2, 3]. The SVM model took bag-of-words TF-IDF features as input, and used chi-square to select top 2,000 features. We applied a linear kernel function for classification.

We evaluated the prediction task using precision, recall and F1. Table 10 shows the performance of our method as well as SVM for divorce prediction.

As shown in Table 10, our cognitive computing framework method (CCF) outperforms SVM by about 6% in F1. Since CCF benefits from the legal semantics understanding layer as well as legal knowledge learning layer, it can recognize informative expressions and better capture the underlying correspondence from fact descriptions to judicial decisions. For example, “The plaintiff and the defendant had a son Cui mou A, and a daughter Cui mou B” and “After marriage they had two children Yang mou A and Yang mou B”, the value of the fact factor variable *childnum* is set to “2” in both cases. There are over 600 legal factors for

Table 7 Result of ownership of custody prediction.

Ownership of custody	Probability(%)
the plaintiff	59.14
the defendant	40.86

Table 9 Result of fee payer prediction.

Fee Payer	Probability(%)
the plaintiff	86.61
the defendant	13.39

Table 10 Performance.

Model	Precision(%)	Recall(%)	F1(%)
CCF	71.22	74.17	72.65
SVM	64.82	68.64	66.67

divorce cases which are very trivial and different through cases, though our work used about the most 50 important of them. Machine learning methods such as SVM is prone to be over-fitting for such a complicated issue.

As shown in the example, along with the prediction decisions, applied induction rules and their corresponding probability are given. It is very helpful for the general public, who are not familiar with professional legal terms, to understand the prediction result, as well as for us to further optimize the system. Compared with other machine learning methods, the output of our model is interpretable.

5 Conclusion

In this paper we provide a cognitive computing system framework to meet the challenges in predicting judicial decisions. Our framework has three layers: legal semantic understanding layer, legal knowledge learning layer and legal knowledge reasoning layer. In our framework, legal factors are represented with a generative grammar and context computing operators, first; Secondly, legal factors are extracted and are used to augment concepts and relations; thirdly a MLN is generated and trained to predict judicial decisions. The experimental results show that, our method can effectively predict the decisions for divorce cases, and offer better performance than SVM; and the predicting decisions are interpretable as applied induction rules are given.

Acknowledgements This work is supported by national high technology research and development plans (863 plan) (No.2013AA064303).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Lin, W.-C., Kuo, T.-T., Chang, T.-J. (2012). Exploiting machine learning models for Chinese legal documents labeling, case classification, and sentencing prediction. *ROCLING XXIV* (pp. 140).
- Liu, Y.-H., Chen, Y.-L., Ho, W.-L. (2015). Predicting associated statutes for legal problems. *Information Processing and Management*, 51(1), 194C211.
- Aletras, N., Tsarapatsanis, D., Preotiuc-Pietro, D., Lamos, V. (2016). Predicting judicial decisions of the European court of human rights: A natural language processing perspective. *PeerJ Computer Science*, 2, e93.
- Luo, B., Feng, Y., Xu, J., Zhang, X., Zhao, D. (2017). Learning to predict charges for criminal cases with legal basis. *Conference on Empirical Methods in Natural Language Processing* (pp. 2727–2736).
- Kelly, I.E. (2015). Computing cognition and the future of knowing. *IBM Research: Cognitive Computing*. IBM Corporation.
- Ludwig, L. (2013). *Extended Artificial Memory. Toward an integral cognitive theory of memory and technology*. Technische University Kaiserslautern.
- Ferrucci, D.A. (2012). Introduction to “this is watson”. *IBM Journal of Research and Development*, 56(3), 1.
- Chen, Y.-L., Liu, Y.-H., Ho, W.-L. (2013). A text mining approach to assist the general public in the retrieval of legal documents. *JASIST*, 64(2), 280C290.
- Raghav, K., Reddy, P.B., Reddy, V.B., Reddy, P.K. (2015). Text and citations based cluster analysis of legal judgments. In *MIKE* (pp. 449C459). Springer.
- Raghav, K., Reddy, P.K., Reddy, V.B. (2016). Analyzing the extraction of relevant legal judgments using paragraph-level and citation information. *AI4JCArtificial Intelligence for Justice* (pp. 30).
- Kim, M.-Y., Xu, Y., Goebel, R. (2014). Legal question answering using ranking svm and syntactic/semantic similarity. In *JSAI International Symposium on Artificial Intelligence* (pp. 244C258). Springer.
- Carvalho, D., Nguyen, M.-T., Tran, C.-X., Nguyen, M.-L. (2016). Lexical-morphological modeling for legal text analysis. *New Frontiers in Artificial Intelligence (Lecture Notes in Artificial Intelligence)*, 10091. <https://doi.org/10.1007/978-3-319-50953-2>.
- Liu, C.-L., & Liao, T.-M. (2005). Classifying criminal charges in chinese for web-based legal services. In *Asia-Pacific Web Conference* (pp. 64C75), Springer.
- Liu, C.-L., Chang, C.-T., Ho, J.-H. (2004). Case instance generation and refinement for case-based criminal summary judgments in chinese. *Journal of Information Science and Engineering*, 20(4), 783C800.
- Liu, C.-L., & Hsieh, C.-D. (2006). Exploring phrase-based classification of judicial documents for criminal charges in chinese. In *International Symposium on Methodologies for Intelligent Systems* (pp. 681C690). Springer.
- Sim, Y., Rouledge, B., Smith, N.A. (2014). The Utility of Text: The Case of Amicus Briefs and the Supreme Court. *Eprint Arxiv*.
- Katz, D.M., Bommarito, I.I., Michael, J., Blackman, J. (2016). A general approach for predicting the behavior of the supreme court of the United States. arXiv:1612.03473.
- Li, J., Li, X., Meng, T. (2015). TML: A General High-performance Text Mining Language[J]. *Computer Research and Development*, 52(3), 553–560.
- Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space[J]. *Computer Science*.
- Zhai, F., Potdar, S., Xiang, B., Zhou, B. (2017). Neural Models for Sequence Chunking. *AAAI*.
- Ma, X., & Hovy, E. (2016). End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF [J]. *ACL*.
- Zeng, D.J., Liu, K., Lai, S.W., Zhou, G., Zhao, J. (2014). Relation classification via convolutional deep neural network. In *The 25th international conference on computational linguistics (COLING2014)* (pp. 2335–2344).
- Singla, P., & Domingos, P. Discriminative training of Markov logic networks. In *National Conference on Artificial Intelligence* (pp. 868–873). AAAI.



Jiaping Li Ph.D The lecturer of China University of Mining and Technology (Beijing). Her main research field is text mining, information extraction, network and information and security.



Longxue Yu Research engineer. She is mainly engaged in text mining, information extraction and computational linguistics.



Guoying Zhang The professor and doctoral tutor of China University of Mining and Technology (Beijing). Her main research field is image recognition and artificial intelligence. She has published more than 50 papers including 40 papers with SCI and EI retrieval and two monograph books in this field. And she won two first and two second in the provincial and ministerial level science and technology progress prizes, obtained 12 the national invention patents.



Tao Meng Ph.D, research scientist. He is mainly engaged in text mining, information extraction and computational linguistics.