



A Review on the Application of Deep Learning in Legal Domain

Neha Bansal^(✉), Arun Sharma, and R. K. Singh

Indira Gandhi Delhi Technical University for Women, Delhi, India
nehabansal33@gmail.com,
{arunsharma, rksingh}@igdtuw.ac.in

Abstract. The Amount of legal information that is being produced on a daily basis in the law courts is increasing enormously and nowadays this information is available in electronic form also. The application of various machine learning and deep learning methods for processing of legal documents has been receiving considerate attention over the last few years. Legal document classification, translation, summarization, contract review, case prediction and information retrieval are some of the tasks that have received concentrated efforts from the research community. In this survey, we have performed a comprehensive study of various deep learning methods applied in the legal domain and classified various legal tasks into three broad categories, viz. legal data search, legal text analytics and legal intelligent interfaces. The proposed study suggests that deep learning models like CNNs, RNNs, LSTM and GRU, and multi-task deep learning models are being used actively to solve wide variety of legal tasks and are giving state-of-the-art performance.

Keywords: Deep learning · Legal text analytics · Classification · Prediction systems

1 Introduction

The continued application of computational intelligence in legal domain has been going on for last few decades. With the increased availability of legal text in digital form, the focus on developing intelligent models and applications have received concentrated rationale from the research community. A wide variety of issues, including summarization, reasoning, classification, translation, text analytics, and others have been applied to a range of legal domain problems. The usage of computer-based intelligent support has many-fold benefits for the legal professional community. These benefits include reducing the laborious human task involved in searching and retrieval of relevant material, reducing the legal costs via automation; resolving or settling issues without the involvement of courts or with less time and effort; negotiation of the law for legal professionals and also the common users; and making decisions based on prediction systems which may be considered more accurate.

In that process, the application of different machine learning and deep learning techniques is crucial. Tasks such as the translation and classification of legal documents, contract reviews as well as the summarization of those are highly relevant.

Deep Learning [1–3] is a specific sub-field of Machine Learning, which is a specific subset of Artificial Intelligence. The concept of deep learning first emerged around 2006. Deep learning is a form of hierarchical learning and involves multiple layers of nonlinear processing for learning high-level abstractions in data [4, 5]. Deep learning is proving to be the next breakthrough in the field of Artificial Intelligence. With state-of-the-art results in solving a wide variety of complex tasks especially related to pattern recognition, image processing and automatic speech recognition, the area promises to hold positive results for further research. Deep learning can be performed as supervised as well as unsupervised learning. The breakthrough in the distributed representation of words using deep learning solidifies the basis of semantic analysis. Many different unsupervised training methods, which generates word embeddings from unstructured data, make the upcoming high-level semantic analysis models achieve the state-of-art results. Deep learning is penetrating its roots in every possible domain and legal domain is also receiving the aforementioned benefits. A lawyer needs to spend hours and hours on searching for relevant material and preparing arguments with relevant precedents. Artificial intelligence enables the human lawyer to work speed and more data. This show us that cooperation of human and AI is important. It aims at providing lawyers more consultancy and getting rid of fatigue duty. This review exclusively covers the recent works employing deep learning models for legal domain and suggests future research directions.

2 Research Methodology

2.1 Literature Selection

We performed an organized review of deep learning works for legal domain. The effective search includes, Journal of Machine Learning (Springer), Journal of Artificial Intelligence and Law (Springer), Nature Scientific Reports, IEEE Conference on Knowledge and Systems Engineering, ACM Conference on Knowledge Discovery and Data Mining, and the International Conference on Artificial Intelligence and Law. We searched using the combinations of keywords from “deep learning,” “neural networks,” “legal data,” “judgments,” and “cases.” We limited our search to recent papers published between January, 2015, and February, 2019, and found total of 78 articles. After going through the title and abstract of all the papers, we limited our study to 14 articles that were studied with full text and further reviewed for the survey.

2.2 Research Questions

With this research we aim to address the following research questions:

- RQ1. What are the available legal datasets to work upon?
- RQ2. What are the activities for legal aspects that have been explored using deep learning? Using this analysis, researchers can identify the best suited deep learning models to work upon a specific legal task.
- RQ3. What are some other activities for legal domain that are still unexplored using deep learning techniques?

3 Literature Review

This section presents a brief discussion on different legal tasks that have been implemented with the help of deep learning models. After reviewing the selected articles, we divided the application of deep learning to legal domain into three broad categories: legal data search, legal text analytics and legal intelligent interfaces as shown in Fig. 1. The first category includes various models developed for retrieving and classifying relevant legal text. The second category includes tasks that require NLP analysis such as summarization, case prediction, identifying sections in legal documents, translation, element extraction from documents. The third category focuses on systems developed to support legal tasks such as question-answering systems, judgment prediction systems and dialogue systems.

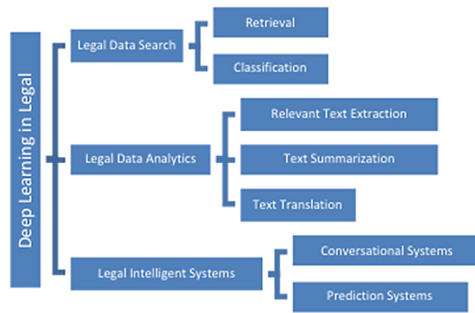


Fig. 1. Categorization of legal activities

3.1 Legal Data Search

A legal domain specific information retrieval system was implemented by Sugathadasa et al. [6]. Authors implemented three different models which incorporated vector space representations of the legal domain. The first model was developed using Node2vec algorithm, second model used sentence similarity and the third was generated using a vector space from both the models and implemented using neural network. Authors concluded that the ensemble model showed higher accuracy level. As further extension, authors concluded that the approach can be used to build information retrieval systems for other domains. Traditional full text search systems finds exact match to a given string and do not take into consideration synonyms and other related terms for each word in the search string. Landthaler et al. [7] worked on an information retrieval system for legal domain that searched for not only the exact matches but also semantically related patterns for any arbitrary length of search query. The system was build using word2vec implementation of word embeddings. As suggested by authors, the system can be further improved by applying various text pre-processing steps such as stemming, stop-word removal, POS tagging and others.

An automated legal document classification model, Supreme Court Classifier (SCC) was implemented by Undavia et al. [8]. Authors compared a number of machine

learning algorithms with the recent NN-based systems. Authors evaluated their system using the Washington University School of Law Supreme Court Database (SCDB). CNN network with word2vec vector performed best and gave an accuracy around 72.4%.

Wei et al. [9] reports preliminary studies in using deep learning for text classification in legal document review. Experiments were conducted on four legal datasets wherein authors compared results of neural network with SVM algorithm. Results showed that CNN gave better accuracy with training dataset of larger size and can be further improved for the text classification in legal industry. A classification system for Brazilian court's document was implemented by Silva [10]. Authors implemented CNN network and obtained satisfactory results.

3.2 Legal Data Analytics

Elnaggar et al. [11] proposed the application of multi-task deep learning model to perform summarization, classification and translation of German legal documents using a single model. Authors suggest that due to the scarcity of German legal documents, a single model was created using the dataset and was used to transfer learning for multiple tasks. Authors concluded that the multi-task Deep learning model outperformed the state of the art results in all three tasks.

A detailed investigation of distributional representations of words and sentences, and the related machine learning and deep learning techniques was done by Wang in his thesis [12]. Author proposed an innovative approach, Word2Sent, for measuring the degree of similarity between sentences. Based on the results, author concluded that the domain-specific word embedding gives better results for the datasets in the domain. An approach based on LSTM model was given by Li et al. [13] for evaluating the rationality of Chinese Judicial decisions. Authors proposed a novel metric, judgment deviation, to measure the likelihood of a certain case's mis-judgment. LSTM model was implemented to extract the elements that effect the decision. Experiments were carried out on Chinese judgments taken from China Judgments online and validation results were satisfactory.

A study on recognizing logical patterns in Vietnamese legal dataset was done by Son et al. [14] using deep learning models. Authors performed experiments using four models based on recurrent neural networks including Long Short Term Memory (LSTM), Bidirectional LSTM and their combination with Conditional Random Fields. Experiments showed that neural networks approaches achieved promising results for this task. Chalkidis et al. [15] developed contract element extraction system using deep learning method. Authors implemented a Bi-LSTM model operating on word, POS tag, and tokenshape embeddings. The system was evaluated using the dataset of 3,500 English contracts having 11 categories of contracts. Authors suggest that by stacking an additional LSTM on top of the Bi-LSTM, or by adding a CRF layer on top of the Bi-LSTM, results were further improved. Authors in their work [16] compared deep learning architectures with traditional algorithms ranging from SVM to ensemble-based decision tree classifiers. Authors present a deep learning architecture for classifying deontic modalities in legal texts. Neural network based classifiers especially LSTM model showed consistent improvement over other classifiers. Authors conclude that further extension is possible by working on other domains.

3.3 Legal Intelligent Interfaces

John et al. [17] worked on a conversational system ‘legalbot’ for legal domain. The system responded to user queries posted as questions. Instead of going for a retrieval based system authors proposed a generative model. The model was build using the Seq2Seq deep learning model. The proposed generative system makes use of domain specific knowledge for generating answers. The system was trained using dataset build from question-answers on some legal concerns. Authors concluded that the results were promising and can be further improved by increasing the dataset provided to the model. Another legal question-answering system was given by Do et al. [18]. The system was build using ranking SVM and convolution neural network. Authors suggest that characteristics of legal text such as references between articles or structured relations in sentences can be explored further to improve the obtained results.

A deep learning based prediction system was proposed by Kowsrihawat et al. [19] for decision of criminal cases. Authors implemented a Bi-directional GRU based decision system for Thai Supreme Court. Earlier systems were build based on bag-of-words model, which generally had a low accuracy as the order of word occurrence is not considered. Recurrent neural networks was implemented to read the fact from an input case and then attention mechanism was used to compare them against relevant legal provisions. The model’s output shows if a person is guilty of a crime or not. The proposed system produced a better F1 score than Naïve Bayes and SVM classification.

Table 1 gives a summary of the legal tasks, approach and the legal dataset on which the approach was validated.

Table 1. Summary of legal tasks, approach and the legal dataset

Legal task category	Sub-category	Approach	Dataset	Results
Legal data search	Document retrieval	Word embeddings with neural network [6]	2500 legal cases from Findlaw website [9-from paper]	NA
		Word embeddings [7]	EU-DPD German rental contracts	NA
	Document classification	Recurrent neural network and convolution neural network [8]	University School of Law Supreme Court Database (SCDB)	word2vec + CNN achieves 72.4%
		Convolution neural networks [9]	Four real legal datasets	CNN outperformed SVM
		Convolution neural network with word embedding [10]	Brazilian court dataset	NA

(continued)

Table 1. (continued)

Legal task category	Sub-category	Approach	Dataset	Results
Legal text analytics	Summarization and translation [11]	Multi-task deep learning model	German Legal Corpora	Multi-task DL performed better than other state-of-the-art works
	Analysing/ extracting relevant text	New word embedding developed [12]	Lawsents (github)	Embedding results were good
		LSTM Model [13]	24,987 Judgments from China Judgments online	Results with artificial validation were accurate
		LSTM, Bi-LSTM, and combination with CRF [14]	Vietnamese legal dataset	Results were satisfactory
		Bi-LSTM model [15]	3,500 English contracts	Results were improved by adding additional LSTM layer or CRF layer
	Non-ANN and ANN models with Legal distributional Semantics model [16]		Results obtained from LSTM model were best	
Intelligent legal interfaces	Conversational agent	Seq2Seq deep learning generative model [17]	Dataset build from question-answer on some legal concerns	As suggested by author, results can be improved further by increasing the dataset size
		Ranking SVM and convolutional neural network [18]	Japan Legal Code	Structured relationship between sentences and references between articles can be explored to improve upon results
	Judgment prediction system	Bi-directional GRU model with attention mechanism [19]	Criminal cases from Thai Supreme court	Results outperformed Naïve Bayes and SVM

4 Conclusion

The use of deep learning and other AI techniques in legal services will accelerate the overall process of judiciary system. The application of deep learning models in various tasks such as legal data search, predictive systems, information retrieval, extraction of relevant text, intelligent interfaces, and legal conversational agents will reduce time, effort and overall cost involved in the domain. From the study, we come to following results:

- Classification of documents is majorly implemented using convolutional neural networks and its variants. Information retrieval systems are enhanced by building domain-specific word embeddings.
- Legal text analytics involving summarization, extraction of relevant text and translation is mostly performed using LSTM models, a variant of recurrent neural network.
- To work on intelligent systems, generative models from deep learning are implemented and providing good results.
- From the datasets, it is also revealed that a number of countries are trying to use deep learning intelligence to improve their judicial systems.

We conclude that the application of deep learning in legal domain has accelerated in last two years, and thus the research is under its initial phase. The comparative evaluation for our survey was not possible as the datasets used in each of the works is unique. The area holds promising future scope, as some other tasks like context-based summarization, predicting the time that will be required to solve a case, and other legal problems can be further explored with the application of suitable deep learning techniques.

References

1. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
2. Soniya, Paul, S., Singh, L.: A review on advances in deep learning. In: *IEEE Workshop on Computational Intelligence: Theories Applications and Future Directions (WCI)*, pp. 14–17. IEEE, India (2015)
3. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
4. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)
5. Bengio, Y.: Learning deep architectures for AI. *Found. Trends@ Mach. Learn.* **2**(1), 1–127 (2009)
6. Sugathadasa, K., et al.: Legal document retrieval using document vector embeddings and deep learning. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) *SAI 2018. AISC*, vol. 857, pp. 160–175. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01177-2_12
7. Landthaler, J., Waltl, B., Holl, P., Matthes, F.: Extending full text search for legal document collections using word embeddings. In: *JURIX*, pp. 73–82 (2016)

8. Undavia, S., Meyers, A., Ortega, J.E.: A comparative study of classifying legal documents with neural networks. In: Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 515–522. IEEE, Poland (2018)
9. Wei, F., Qin, H., Ye, S., Zhao, H.: Empirical study of deep learning for text classification in legal document review. In: IEEE International Conference on Big Data (Big Data), pp. 3317–3320. IEEE, USA (2018)
10. Da Silva, N.C.: Document type classification for Brazil’s supreme court using a convolutional neural network. In: The Tenth International Conference on Forensic Computer Science and Cyber Law-ICoFCS, pp. 7–11. Brazil (2018)
11. Elnaggar, A., Gebendorfer, C., Glaser, I., Matthes, F.: Multi-task deep learning for legal document translation, summarization and multi-label classification. arXiv preprint [arXiv: 1810.07513](https://arxiv.org/abs/1810.07513) (2018)
12. Wang, Y.: An unsupervised approach to relatedness analysis of legal language. Master’s thesis, University of Waterloo (2018)
13. Li, S., Zhang, H., Ye, L., Guo, X., Fang, B.: Evaluating the rationality of judicial decision with LSTM-based case modeling. In: IEEE Third International Conference on Data Science in Cyberspace (DSC), pp. 392–397. IEEE, China (2018)
14. Son, N.T., Nguyen, L.M., Quoc, H.B., Shimazu, A.: Recognizing logical parts in legal texts using neural architectures. In: Eighth International Conference on Knowledge and Systems Engineering (KSE), pp. 252–257. IEEE, Vietnam (2016)
15. Chalkidis, I., Androutsopoulos, I.: A deep learning approach to contract element extraction. In: JURIX, pp. 155–164 (2017)
16. Neill, J.O., Buitelaar, P., Robin, C., Brien, L.O.: Classifying sentential modality in legal language: a use case in financial regulations, acts and directives. In: Proceedings of the 16th Edition of the International Conference on Artificial Intelligence and Law, pp. 159–168. ACM, USA (2017)
17. John, A.K., Di Caro, L., Robaldo, L., Boella, G.: Legalbot: a deep learning-based conversational agent in the legal domain. In: Frasinca, F., Ittoo, A., Nguyen, L.M., Métais, E. (eds.) NLDB 2017. LNCS, vol. 10260, pp. 267–273. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59569-6_32
18. Do, P.K., Nguyen, H.T., Tran, C.X., Nguyen, M.T., Nguyen, M.L.: Legal question answering using ranking SVM and deep convolutional neural network. arXiv preprint [arXiv: 1703.05320](https://arxiv.org/abs/1703.05320) (2017)
19. Kowsrihawatt, K., Vateekul, P., Boonkwan, P.: Predicting judicial decisions of criminal cases from thai supreme court using bi-directional GRU with attention mechanism. In: 5th Asian Conference on Defense Technology (ACDT), pp. 50–55. IEEE, Vietnam (2018)