

DLRG@AILA 2019:Context - Aware Legal Assistance System

R.Rameshkannan and R. Rajalakshmi *

School of Computing Science and Engineering
Vellore Institute of Technology, Chennai, India
ramehskannan.r@vit.ac.in, rajalakshmi.r@vit.ac.in

Abstract. In this digital era, seamless information is available in the web. The Information Retrieval systems play an important role in quickly retrieving the relevant information based on the query from the user. Common Law Systems are followed in countries like UK, USA, Canada, Australia and India that has two primary sources of law viz., statutes (established laws) and precedents (prior cases). The statutes deal with applying legal principles to a situations which may lead to filing the case, and the precedents help lawyers to understand how the Court has handled the similar scenarios in the past, for the subsequent legal reasoning. For any given situation, applying the appropriate statutes as well as identifying the prior cases are important and it is a time consuming process. There is a demand for an automated system which can identify the set prior cases and suitable statutes for any situation. This will help the lawyers to get a preliminary understanding of the case and to identify where the problem fits. The objective of this work is to develop such an automatic system to identify relevant law or prior cases for a given situation. This work has been submitted to AILA 2019 (Artificial Intelligence for Legal Assistance). Here, the assigned task is to identify the relevant statute (task1) / prior cases (task2) for a given situation, by considering the Indian Legal documents. For this legal document retrieval task, we present our context-aware solution that finds the similarity between the given situation and legal documents / prior cases by following an effective word representation that considers the dependency between the terms. We have evaluated our methodology on the dataset released by the organizers of AILA@FIRE2019 shared task. We have used the p@10 score as the evaluation metric, and achieved the score 0.015 and 0.05 for task1 and task2 respectively.

Keywords: Legal Document Retrieval, TF-IDF, Glove, Word representation, Information Retrieval

* Corresponding Author

1 Introduction

Many of the information systems are being developed in the current world of information society to encourage users to make effective use of information collection. Over the past few decades, there is a significant evolution in information systems. But the rate of development is very slow and finding difficulty in relating documents each other. Lawyers' time is wasted on judicial process for finding the relevant document and understanding the documents. So there is a need for automated systems to be used by the lawyers for legal purpose. When a lawyer is presented with a situation that may lead to a case being filed, it will be very beneficial to him / her if there is an automatic system that identifies a set of related preliminary cases involving similar situations as well as statutes or acts that may be most suitable for the purpose in the given situation. Such a system not only helps a lawyer, but also benefits a common man, in order to gain a preliminary understanding even before he / she approaches a lawyer. It shall assist him / her in identifying where his / her legal problem fits, what legal actions he / she may take through statutes, and what the outcomes of similar cases have been through precedents. Finding the text similarity or relevance is one of the important task in information systems to retrieve the relevant documents. It is also used in multiple tasks such as information ranking, recognition of paraphrases and plagiarism detection.

In this work, we describe our method that was submitted to the AILA 2019 (Artificial Intelligence for Legal Assistance) shared task. As part of the data challenge, two tasks have been assigned. Task 1 is to identify the relevant prior case and the Task 2 is to identify the most relevant statutes for any given situation / query. This task can be viewed as a unsupervised problem and text similarity between the query and the legal documents can be considered for effective retrieval. For both the tasks, only Indian legal documents have been considered. with Indian statutes and prior cases decided by Indian courts of Law. The relevance between the query and the document is calculated by using two different word representation viz., TF-IDF and Glove to find the similarity between the words in the query and the legal documents / prior cases. From the experimental results on the dataset, we observed that, considering the word dependency is useful for the retrieval of relevant documents. We have used the p@k measure as the evaluation metric and achieved a score of 0.015 and 0.05 for the task1 and task2 respectively. The methodology is presented in the following section along with the results and discussion.

2 Materials and Methods

In recent years, the recommendation systems have gained popularity in various domains. In this legal assistance recommendation systems, the judgments or statutes of older cases play a vital role in terms of recommendations. To obtain suitable recommendations, the appropriate similarity measure should be chosen that can find similar prior cases or the relevant laws that fit in correctly

for the given situation. Different measures have been reported in the existing literature on recommendation systems in various domains. Lucas Colucci et.al. [1], suggested TF-IDF for feature extraction and cosine similarity for calculating the relatedness of the movies by using the TMDb dataset. Hongkun Leng ,et al [2],proposed recommended system with TF-IDF,BM25F and Jaccard similarity approach and shown that TF-IDF feature weighting method resulted in high Mean Average Precision score. In information retrieval tasks, many different feature weighting methods have been used. Among the various techniques, TF-IDF is found to be effective for text categorization [7]. Other techniques include CHI square [8],[5] and mutual information based feature selection [6]. Also, the effectiveness of such techniques along with SVM and CNN have been studied [9], [10].

2.1 Dataset Description

In this work, we have carried out the experiments using the dataset released by the organizers of AILA 2019 task [3]. The objective of the Task1 is to identify the relevant prior cases for the given situation. To perform Task1, a set of 50 queries (Q1 to Q50) that contain the description of legal situations along with 2914 labled prior cases (C1 to C2914) documents have been used. Among the released 2914 total prior cases, the task is to identify those cases that are relevant to the queries. For performing Task2, 197 statues were released that contain title and textual description and the task is to retrieve the most relevant statues using the same queries (Q1 to Q50).

2.2 Methodology

For both the tasks, identifying the relevant prior cases and statues from the queries, the queries from 1 to 50 were taken. All the queries which needs to be used for the process are given as input to the system it may be either cases/statutes and query. All the inputs are preprocessed and tokenization was done by utilizing NLTK[4] and then stop words were removed. In preprocessing stage, only the alphabets are considered and all the other characters are removed. The traditional TF-IDF feature weighting method has been applied and the similarity between the query and case docuemnts were calculated. In this approach, the words are represented in a sparse representation and dependancy was not considered. Next, we have used the cosine distance as the similarity measure and performed the second experiment. In the third experiment, we have applied a dense vector representation using GLOVE and considered the dependency between the terms using their sequence information.

For the AILA Task1 and Task2 ,the obtained results are tabled using the evaluation methods like p@10,Mean Average Precision(MAP).

Precision at k (P@k): Precision at k is the proportion of recommended items in the top-k set that are relevant. It is the mean of the precision calculated over all the topics of the first ten(k) documents retrieved.

$$f(x) = \frac{\text{Number of recommended items } k \text{ that are relevant}}{\text{Number of recommended Items @}k} \quad (1)$$

Mean average precision (MAP): Q is the number of queries in the set and $AveP(q)$ is the average precision (AP) for the given query(q).

$$f(x) = \frac{\sum_{q=1}^Q AveP(q)}{Q} \quad (2)$$

For a given query(q), we need to calculate its corresponding Average Precision, and then the mean of the all these Average Precision scores would give us a single number, called the mean Average Precision. The mean Average Precision (MAP) value shows how our system is performing for the given query(q).

The first approach Cosine distance uses the angle representations of the words to find the related documents. From table1 and table2, the result attained are using the cosine distance approach was 0.0075 for task1 and 0.0225 for task2. The result computed with p@10 result is not enough to attain good result. So we tried with tf-idf for legal relationship with the document and query. Here the related and non related words are taken into account for performance. term frequency used to find the relevance between the document. The score attained with 0.1 for task1 and 0.035 for task2. And the final approach Glove performs with Dense representational format. Using glove representation, cosine distance distance is calculated for each and every word from the document. The result achieved by using glove is 0.015 for task1 and 0.05 for task2.

Table 1. Case/ Precedent Results

S.No	Case Doc	p@10	map
1	Cosine distance	0.0075	0.009
2	Tf-idf Cosine	0.01	0.0416
3	Glove	0.015	0.0432

Table 2. Statute Results

S.No	Statute Doc	p@10	map
1	Cosine distance	0.0225	0.03
2	Tf-idf Cosine	0.035	0.06
3	Glove	0.05	0.089

For in case of, prior cases situation the same set of experiments are conducted. But the attained results had good effect on Glove method compared to other two methods namely tf-idf and cosine similarity ,with a p@10 score of 0.015 and 0.05 for task1 and task2 respectively.

For the task1, Identifying relevant prior cases achieved a p@10 score of 0.015 and for the task 2, Identifying relevant statutes achieved a p@10 score of 0.05 by utilizing Glove vector representations. For calculating similarity, the minimum distance between the vectors are taken into account for every term in query and the document. For each of the information retrieval, the set of submitted task runs be ordered from the highest value in the metric to the lowest in the same metric. That is best performing metric to the worst performing metric. We compared the obtained orderings between different metrics using p@10,MAP.

Furthermore, the information retrieval metrics considered in this paper are all based on either document is relevant to the query or not.

3 Conclusion

In this paper, we described about the feature weighing method for the recommendation system .We studied cases and statutes of previous judgments to recommend what they want in response to the queries they had to get suggestions. We have tried several contributions to the recommended system. We have also described about the different methods used to evaluate the queries. From the above implementations, we achieved for task1 p@10 score of 0.015 and for the task2 with Glove representation attained p@10 score of 0.05. The system performance can be improved by applying different features and different similarity measures.

4 Acknowledgment

The authors would like to thank the management of Vellore Institute of Technology (VIT), Chennai for providing the support to carry out this work. We would also like to thank the Department of Science and Engineering Research Board (SERB), Government of India for their financial grant (Grant awarded ECR/2016/00484) for this research work.

References

1. Colucci, Lucas and Doshi, Prachi and Lee, Kun-Lin and Liang, Jiajie and Lin, Yin and Vashishtha, Ishan and Zhang, Jia and Jude, Alvin.: Evaluating Item-Item Similarity Algorithms for Movies. In:2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp.2141–2147,ACM, New York, NY, USA (2016)
2. Hongkun Leng and Caleb De La Cruz Paulino and M. Haider and Rui Lu and Zhenhui Zhou and Ole J. Mengshoel and Per-Erik Brodin and Julien Forgeat and Alvin Jude.: Finding Similar Movies: Dataset, Tools, and Methods.(2018)
3. P. Bhattacharya, K. Ghosh, S. Ghosh, A. Pal, P. Mehta, A. Bhattacharya., P. Majumder, Overview of the Fire 2019 AILA track: Artificial Intelligence for Legal Assistance. In Proc. of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, (2019).
4. Edward Loper and Steven Bird,NLTK: The Natural Language Toolkit.In:ACL workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. Philadelphia: Association for Computational Linguistics,(2002)
5. Rajalakshmi, R., Agrawal, R., Borrowing Likelihood Ranking based on Relevance Factor, In: Proceedings of the Fourth ACM IKDD Conferences on Data Sciences, CODS 2017, India, pp: 12:1–12:2
6. Rajalakshmi, R., Xavier, S., Experimental Study of Feature Weighting Techniques for URL Based Webpage Classification, Procedia Computer Science, Vol.115, pp. 218-225, (2017)

7. Sivakumar, S., Rajalakshmi, R, Comparative evaluation of various feature weighting methods on movie reviews, *Advances in Intelligent Systems and Computing*, Vol-711, pp. 721-730 (2019).
8. Rajalakshmi, R., Aravindan, C., Naive Bayes approach for URL classification with supervised feature selection and rejection framework, *Computational Intelligence*, 34(1), pp. 363-396 (2018).
9. R. Rajalakshmi, C. Aravindan, "An Effective and Discriminative Feature Learning for URL Based Web Page Classification," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 1374-1379, (2018).
10. Rajalakshmi, R., Ramraj, S., Ramesh Kannan, R, Transfer learning approach for identification of malicious domain names, *Communications in Computer and Information Science*, Vol. 969, pp. 656-666.(2019)