

Legal Information Retrieval and Rhetorical Role Labelling for Legal Judgements

Nitin Nikamanth Appiah Balaji, B. Bharathi and J. Bhuvana

Department of CSE, Sri Siva Subramaniya Nadar College of Engineering, Tamil Nadu, India

Abstract

Retrieving the most relevant information from a huge amount of documents is a tedious process. Nowadays, artificial intelligence plays a major role in information retrieval tasks. Information retrieval tasks can be used in different applications like search engines, relevance feedback, summarization, and so on. In this proposed work explains the method used to solve the problem given in the shared task on Artificial Intelligence for Legal Assistance proposed by the Forum of Information Retrieval Evaluation in 2020(AILA@Fire2020). There are two tasks given in the challenge. The first task is, given the description of a situation, identify relevant statutes, and prior-cases. The second task is, given a legal case document, classify each sentence of the document in one of the 7 semantic segments/rhetorical roles. We compare two systems for the first task - TFIDF features along with the cosine similarity metric and BM25 ranking algorithm. For the semantic labeling task, pre-trained FastText embedding with MLP and TFIDF with random forest classifiers are used. BM25 ranking algorithm shows significantly better results for the first task and the pre-trained FastText method is better than the vectorization method for classifying the rhetorical roles.

Keywords

Legal information retrieval, Rhetorical Role labelling, BM25, TFIDF, FastText

1. Introduction

The countries like India, UK, Canada, Australia, and many others use the common law system. There are two important primary sources of law that exists in the system. The first one is called statutes which are the written laws. The precedents or judgments of prior cases delivered by a court, which involve similar legal facts and issues are the current case but are not directly indicated in the written law. When the legal counselor working on a new case often relies on these statutes and precedents to understand how the Court has discussed, argued, and behaved in similar scenarios. The first task is to retrieve the relevant statutes and prior-cases by giving the situation of the problem.

Most of the legal case documents follow a common structure with different sections like "Details of the Case", "Issues being discussed", "Arguments given by the parties", etc. These sections are popularly termed as "rhetorical roles". An acquaintance of such semantic roles will not only improve the readability of the documents but also needed to compute document


FIRE 2020: Forum for Information Retrieval Evaluation, December 16-20, 2020, Hyderabad, India

✉ nitinnikamanth17099@cse.ssn.edu.in (N. N. A. Balaji); bharathib@ssn.edu.in (B. Bharathi); bhuvanaj@ssn.edu.in (J. Bhuvana)

ORCID 0000-0002-6105-0998 (N. N. A. Balaji); 0000-0001-7279-5357 (B. Bharathi)



© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

similarity, summarization, etc. However, this information is generally not specified explicitly in case documents, which are usually just free-flowing text. The second task is to semantically label the sentence into one among the seven roles.

2. Related work

In AILA 2019, task of identifying most relevant prior cases for the given situation and retrieving most relevant statutes for the given situation was solved by different authors. In [1], BM25 ranking algorithm with Doc2vec unsupervised algorithm is used for the task. The named entity recognition preprocessing with TFIDF and BM25 was used in [2]. In [3], pre-trained word embeddings are used for the query and relevant documents. Cosine similarity was used for retrieving the most relevant prior cases or statutes. The authors of [4] used different vectorization methods with similarity metrics such as jaccard similarity and cosine similarity to retrieve the target documents. In [5], Language Model for Information Retrieval, Vector Space Model and BM25 were used for information retrieval. In [6], the important topic words were extracted from the given situation and the topic words were used as a query to identify the relevant prior cases.

3. Proposed system

3.1. Precedent & Statute Retrieval

The overview of the task is described in [7]. The first task involves the identification and ranking of the most relevant statutes or prior-cases for a given description of a legal situation. This is done by checking the correlation between the requested query and the available prior documents. Based on the correlation based on the algorithm used the documents are ranked. For this task, BM25 ranking algorithm and TFIDF vectorization along with cosine similarity are considered. The performance of both the systems is compared with MAP, bpref, Recip_rank, P@10 scores.

3.1.1. TFIDF - Cosine Similarity

TFIDF (Term Frequency Inverse Document Frequency) is used to convert the given queries and the documents into number vectors. TFIDF generates a vector removing the significance of stop words and words with less semantic meaning. Reducing this noise from the documents can help the cosine-similarity function to concentrate on the important words. The cosine-similarity produces the correlation between the query and the document by the equation 1.

$$\cos \theta = \frac{Q \cdot D}{|Q| \cdot |D|} \quad (1)$$

where Q is the TFIDF vector of the query and D is the TFIDF vector of the document - statute or case.

3.1.2. BM25 Ranking

BM25 (Best Matching) algorithm is a bag of words based algorithm which ranks based on the appearance of query key terms on each document. It does not consider the proximity of the query keys in the documents. The BM25 is between a query and a document is calculated using the equation 2. The Okapi BM25 algorithm is used for this task.

$$score(Q, D) = \sum_{i=1}^n IDF(q_i) \frac{f(q_i, D)(k_1 + 1)}{f(q_i, D) + k_1(1 - b + b \frac{|D|}{avgdl})} \quad (2)$$

where q_1, \dots, q_n are the keywords in Q and D is the Document. $f(q_i, D)$ is q_i 's term frequency. $|D|$ is the number of words in D . $avgdl$ is the average length of documents. $IDF(q_i)$ is the inverse document frequency. k_1 and b are free parameters.

3.2. Rhetorical Role Labeling for Legal Judgements

The second task involves the classification of sentences from legal case documents into 7 semantic segments or rhetorical roles. For converting the sentences into a numerical feature matrix two different feature extraction techniques - FastText and TFIDF are implemented and compared. The features thus extracted are classified for the role labels by using Multilayer Perceptron (MLP) and Random Forest (RF) classifiers. The scikit-learn implementation of the machine learning models and the TFIDF text feature extractor is used. The accuracy, macro F1, precision, and recall values are used for evaluation.

3.2.1. FastText Embedding

FastText pre-trained models are CBOW models trained on CommonCrawl and Wikipedia data. As the number of data samples is limited, the pre-trained embeddings could provide scope for improvement. The FastText model pre-trained on the English corpus is considered. The embedding of sentences is extracted for the sentences in the data-set. 300 dimensions fixed-length vector is generated by each sentence in the data-set. This is then fed to a Multilayer Perceptron with 512, 128 hidden layer size, and trained for 200 iterations.

3.2.2. TFIDF Vectorization

TFIDF vectorization is based on the count-based vectorization technique. In TFIDF, the inverse document frequency term reduces the impact of the common words on the classification task. TFIDF vectorization with an n-gram range of 1-5 is taken and a sparse matrix is generated. This sparse matrix is then classified using a Random Forest classifier.

4. Results and Discussions

On comparing the P@10 for the precedent records retrieval task, it is clear that the BM25 algorithm outperforms the TFIDF model. But the statute retrieval task, the TFIDF model

Table 1

Results of test-set for Task1 - Precedent & Statute Retrieval.

Task	Model	MAP	BPREF	Recip_Rank	P@10
Precedent	BM25	0.1264	0.0918	0.2043	0.08
	TFIDF + cosine-similarity	0.0652	0.0406	0.1004	0.05
Statute	BM25	0.1181	0.069	0.2739	0.07
	TFIDF + cosine-similarity	0.3423	0.136	0.3423	0.07

Table 2

Results of test-set for Task2 - Rhetorical Role Labeling for Legal Judgements.

Model	Classifier	Precision	Recall	Macro F1 score	Accuracy
FastText	MLP	0.384	0.4	0.354	0.46
TFIDF n-gram=1-5	RF	0.473	0.354	0.333	0.467

performance slightly better. The P@10 score of both the models is equal for the statute subtask. So we can say that the BM25 model performed better than the TFIDF model. For task2, the FastText and the TFIDF models produce models with similar performance. The accuracy of both the models is 0.46%, but when comparing the macro F1 score, FastText shows slightly better performance, around 6% improvement over the TFIDF model. This is because of the advantage the FastText model gets from the pre-trained weights. The performance of the models on test-data are shown in Table 1 and 2 for task1 and task2 respectively.

5. Conclusion

In this paper, we study the methods for legal document retrieval and rhetorical role labeling. We see that information retrieval techniques play a major role in these tasks. We propose BM25 and TFIDF - cosine similarity algorithms for the precedent and statute retrieval task. Our model achieved a P@10 score of 0.08 and 0.07 for precedent and statute retrieval sub-tasks. For the rhetorical role labeling task, we compare the FastText and TFIDF models. Our FastText model produces a macro F1 score of 0.354.

References

- [1] B. Gain, D. Bandyopadhyay, A. De, T. Saikh, A. Ekbal, IITP at AILA 2019: System report for artificial intelligence for legal assistance shared task, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019, volume 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 19–24. URL: <http://ceur-ws.org/Vol-2517/T1-3.pdf>.
- [2] R. More, J. Patil, A. Palaskar, A. Pawde, Removing named entities to find precedent legal cases, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019, volume

- 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 13–18. URL: <http://ceur-ws.org/Vol-2517/T1-2.pdf>.
- [3] S. Mandal, S. D. Das, Unsupervised identification of relevant cases & statutes using word embeddings, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation*, Kolkata, India, December 12-15, 2019, volume 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 31–35. URL: <http://ceur-ws.org/Vol-2517/T1-5.pdf>.
- [4] S. Kayalvizhi, D. Thenmozhi, C. Aravindan, Legal assistance using word embeddings, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation*, Kolkata, India, December 12-15, 2019, volume 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 36–39. URL: <http://ceur-ws.org/Vol-2517/T1-6.pdf>.
- [5] Y. Shao, Z. Ye, Thuir@aila 2019: Information retrieval approaches for identifying relevant precedents and statutes, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation*, Kolkata, India, December 12-15, 2019, volume 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 46–51. URL: <http://ceur-ws.org/Vol-2517/T1-8.pdf>.
- [6] Z. Zhao, H. Ning, L. Liu, C. Huang, L. Kong, Y. Han, Z. Han, Fire2019@aila: Legal information retrieval using improved BM25, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation*, Kolkata, India, December 12-15, 2019, volume 2517 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 40–45. URL: <http://ceur-ws.org/Vol-2517/T1-7.pdf>.
- [7] P. Bhattacharya, P. Mehta, K. Ghosh, S. Ghosh, A. Pal, A. Bhattacharya, P. Majumder, Overview of the FIRE 2020 AILA track: Artificial Intelligence for Legal Assistance, in: *Proceedings of FIRE 2020 - Forum for Information Retrieval Evaluation*, 2020.