Overview of the Third Shared Task on Artificial Intelligence for Legal Assistance at FIRE 2021

Vedant Parikh[a,b], Upal Bhattacharya[c], Parth Mehta[d], Ayan Bandyopadhyay[e], Paheli Bhattacharya[f], Kripabandhu Ghosh[c], Saptarshi Ghosh[f], Arindam Pal[g], Arnab Bhattacharya[h] and Prasenjit Majumder[a,e]

[a]*Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India*
[b]*Amazon, India*
[c]*Indian Institute of Science Education and Research (IISER), Kolkata, India*
[d]*Parmonic AI, USA*
[e]*TCG Crest, India*
[f]*Indian Institute of Technology Kharagpur, India*
[c]*Indian Institute of Science Education and Research (IISER), Kolkata, India*
[g]*Data61, CSIRO and UNSW, Sydney, New South Wales, Australia*
[h]*Indian Institute of Technology Kanpur, India*

Abstract

The third edition of the shared task on Artificial Intelligence for Legal Assistance (AILA 2021) focused on two problems: (1) Rhetorical Role labeling for Legal Judgements and (2) Legal Document Summarization. Task 1 was a continuation from the previous year, where the objective is to assign one of the rhetorical labels – Facts of the case, Ruling by the Lower Court, Argument, Statute, Precedent, Ratio of the decision, and Ruling by the Present Court – to each sentence in a case judgement (of the Supreme Court of India). For Task 2, entire texts of Supreme Court judgements were provided and the task was to generate a summary by selecting the most important content. Task 2 was further divided into two sub-tasks: (2a) Identifying "summary-worthy" sentences in a court judgement, and (2b) Generating a summary from a given court judgement.

Keywords

Rhetorical role labeling, Semantic segmentation, Legal document summarization, Headnote generation

1. Introduction

With the recent advances in Natural Language Processing and Machine Learning there is a renewed interest in utilizing these techniques to achieve a better understanding of legal

FIRE 2021: Forum for Information Retrieval Evaluation, December 13-17, 2021, India
vedant.parikh.6299@gmail.com (V. Parikh); upal.bhattacharya@gmail.com (U. Bhattacharya); parth.mehta126@gmail.com (P. Mehta); bandyopadhyay.ayan@gmail.com (A. Bandyopadhyay); paheli@cse.iitkgp.ac.in (P. Bhattacharya); kripa.ghosh@gmail.com (K. Ghosh); saptarshi@cse.iitkgp.ac.in> (S. Ghosh); arindamp@gmail.com (A. Pal); arnabb@cse.iitk.ac.in (A. Bhattacharya); prasenjit.majumdar@gmail.com (P. Majumder)

documents. The series of AILA (Artificial Intelligence for Legal Assistance) shared tasks is aimed at solving some of the pressing problems in this domain [1, 2, 3].

AILA 2021 had two tasks. Task 1 is based on the premise that Legal case documents follow a common thematic structure with implicit sections like 'Facts of the Case', 'Issues being discussed', 'Arguments given by the parties', etc. popularly termed as "rhetorical roles" [4]. Knowledge of such semantic segments or roles will not only enhance the readability of the documents but also help in downstream tasks like computing document similarity, summarization [5], etc. However this information is generally not specified explicitly in case documents, which are usually just very long and unstructured, lacking section and paragraph headings. AILA 2021 - Task 1 is a continuation from the previous edition of AILA[2] and is aimed at addressing this gap.

A new task on Legal document Summarization was introduced this year. Indian Judiciary is one of the largest Judicial Systems in the world, consisting of the Supreme Court of India, 25 High courts and 72 District courts, all of which produce an enormous amount of data in form of judgements and other documents. Some of these judgements can run into hundreds of pages. Legal practitioners generally depend on manually written summaries, also known as *Headnotes*, while referring to these judgements. However creating Headnotes takes considerable human effort and is a very slow process. There is an immediate need for automatically creating these Headnotes. Task 2 is aimed at addressing this gap.

AILA 2021 witnessed a participation of 15 teams with 9 of them submitting the working notes. We received a total of 74 runs across all the tasks. The participating teams are distributed across India, China, Botswana, Italy, Austria and Canada, both from academic institutions as well as industry.

1.1. Task 1 : Rhetorical Role Labeling for Legal Judgements

A case document from the Indian judiciary is usually very long and unstructured, without any section / paragraph headings. Therefore, knowledge of which sentence belongs to which particular rhetorical class may enhance process of understanding the document. To this end, this task aims to classify each sentence of the document in one of the 7 semantic segments/rhetorical roles explained below:

- Facts: sentences that denote the chronology of events that led to filing the case
- Ruling by Lower Court : since we deal with Indian Supreme Court cases, these cases were given a preliminary ruling by some lower courts (Tribunal, High Court etc.). These sentences correspond to the ruling/decision given by these lower courts.
- Argument: sentences that denote the arguments of the contending parties
- Statute: relevant statute cited is the ongoing case
- Precedent: relevant precedent (prior case) cited in the ongoing case
- Ratio of the decision : sentences that denote the rationale/reasoning given by the Supreme Court for the final judgement

- Ruling by Present Court : sentences that denote the final decision given by the Supreme Court for the ongoing case

1.2. Task 2 : Summarization of Legal Judgements

In a common law system like in India, past verdicts delivered by various courts in the country are very important. These can be used as a legal basis for arguing future cases. As such, there is a lot of interest in analyzing the past verdicts. Number of such judgements are increasing at an exponential rate which makes it difficult for any legal professional to analyze each case in detail. Instead they often rely on human written summaries a.k.a. 'Headnotes' which are generated by Law professionals. However, getting headnotes written manually by law professionals is a slow and expensive process, and there is a huge interest in the legal community to generate these summaries automatically. However, few such systems exist to date for the Indian legal documents [5], which pose a challenge of their own due to lack of a defined structure. Acknowledging this gap, we include a legal document summarization task in AILA this year. The aim is to automatically generate a shorter version of original document that represents the most important or relevant information. The task was further divided into two sub-tasks as described below.

1.2.1. Task 2a : Identifying 'summary-worthy' sentences in a court judgement

Given a judgement the task is to identify sentences which are 'summary worthy', i.e. which have at least some information that should be included in the summary. Often the facts and rationale of the judgement are given more importance compared to a precedence while creating a Headnote. In Task 2a, we aim to replicate the behaviour. We appreciate the fact that generating / compressing sentences is not an easy task, but a system that can at least identify the "interesting" source sentences can still be of help to legal professionals. This task can be seen as a sentence classification task where each sentence can be either "summary worthy" or not.

1.2.2. Task 2b : Automatically generating a summary from a given court judgement

Task 2b builds on Task 2a, and aims to generate an actual summary as opposed to just selecting informative sentences. This can be seen as a more abstractive summarization as opposed to the extractive nature of Task 2a. Given a court judgement, the participants have to automatically generate a summary for it. This subtask can be seen as a continuation of Task 2a or as a separate subtask. For instance, the summary could simply be formed by collecting and reordering the sentences identified as important in Task 2a. On the other hand, these sentences can be compressed/re-written, or generative models can be used to obtain summaries of an abstractive nature.

2. Dataset

For both tasks we annotate publicly available judgements delivered by the Supreme Court of India. The details of annotations for task 1 can be found in [4] and those for task 2 in [6].

Task 1: The *training set* consisted of 60 annotated documents containing approx. 11,300 sentences in total across all the documents. These consist of the combined training and test set from AILA 2020 [2]. The rhetorical labels were assigned by law experts from a reputed law school in India. As the *test set*, we consider a set of 10 additional case documents (2 documents from each of the 5 law domains mentioned in [4]). These documents were then given to a law expert for annotating every sentence with one of the rhetorical labels. There are a total of 850 sentences in the test set. A part of this dataset is also made publicly available by [4].

Task 2: The training data for Task 2 consisted of 500 document-summary pairs. Each judgement is accompanied by a summary written by a legal expert. Since regular NLP pipeline (stemmer, tokenizers, etc) doesn't work well with legal documents, we provide a pre-processed and sentence tokenized version of each document and summary. For each sentence in the judgement text we will provide a noisy label (75% accurate), which indicates whether or not the sentence is 'summary-worthy'. Each judgement and summary sentence are additionally labelled with one of the seven rhetorical roles mentioned in Task 1. The 'summary-worthy' label as well as the rhetorical roles are assigned automatically and are noisy. Details of this annotation is available in [6]. In total, the training data contained 72,192 sentences. The test data was of 50 document-summary pairs manually annotated with summary worthy labels. This contained a total of 5,066 sentences.

For Task 2, since ROUGE scores are sensitive to the length of auto generated summary, which can vary drastically across judgements, we provide the target summary lengths beforehand. Participants are expected to generate summaries of length as close as possible to the specified target length.

3. Evaluation

For Task 1, evaluation methodology was same as that in AILA 2020 [2]. Standard metrics of Recall, Precision and F1-Scores were used to rank the systems. Since there is a class imbalance among the 7 categories / rhetorical roles, we use macro-averaging at category-level. The scores were calculated as below:

1. Recall, Precision and F-score were computed for each category of labels across all documents.
2. The overall scores for a run are computed by averaging the scores across all categories.

For Task 2a, we use the standard classification metrics Precision, Recall and F1-Score. We further also report the accuracy, which is percentage of labels predicted correctly. All these metrics are averaged across all documents.

For Task 2b, we use the standard ROUGE metrics for ranking the submissions. We specifically use ROUGE-1, ROUGE-2 and ROUGE-4 metrics for this. As, ROUGE Scores are sensitive to document length, for each judgement ideal length was provided.

Team Name	Run ID	P	R	F	Method Summary
Rustic	1	**0.548**	0.616	**0.557**	BiGRU + CRF with Domain specific embeddings
Rustic	2	<u>0.528</u>	<u>0.619</u>	<u>0.551</u>	BiGRU + CRF with Domain specific embeddings
Rustic	3	0.511	**0.627**	0.549	BiGRU + CRF with Domain specific embeddings
MiniTrue	1	0.485	0.572	0.517	Legal Bert + simple neural inference network
Arguably	1	0.465	0.591	0.505	Ernie 2.0
MiniTrue	3	0.461	0.570	0.503	Roberta + simple neural inference network
MiniTrue	2	0.460	0.565	0.501	BigBird + simple neural inference network
SSN_NLP	2	0.451	0.571	0.491	BERT
Arguably	2	0.45	0.586	0.491	Ernie 2.0 with preprocessed sentences
SSN_NLP	3	0.438	0.571	0.475	Roberta
NITS Legal	2	0.453	0.464	0.451	LegalBERT and SMOTE for oversampling
NITS Legal	1	0.441	0.434	0.428	LegalBERT and MLP
SSN_NLP	1	0.411	0.539	0.409	LaBSE
Legal AI 2021	1	0.394	0.361	0.364	MiniLM and SVM
UB_BW	3	0.336	0.369	0.340	Fastext Classifier with Trigram Features
UB_BW	2	0.335	0.371	0.338	Fastext Classifier with Bigram Features
Chandigarh Concordia	3	0.317	0.488	0.329	Sugeno Integral using finetuned BERT
Chandigarh Concordia	2	0.317	0.485	0.327	BERT
UB_BW	1	0.301	0.366	0.317	Fastext Classifier with Unigram Features
Chandigarh Concordia	1	0.29	0.476	0.298	BERT
Legal NLP	3	0.225	0.227	0.22	BERT
CEN NLP	2	0.309	0.19	0.199	DistilBERT
Legal NLP	1	0.197	0.217	0.196	Cased BERT
Legal NLP	2	0.198	0.215	0.192	BERT
CEN NLP	1	0.179	0.194	0.179	DistilBERT
Nit Agartala	1	0.192	0.22	0.179	BERT with preprocessing techniques

Table 1
Results of Task 1: Rhetorical Role Labeling for Legal Judgements. Measures averaged over 10 test documents. Numbers in **bold** and <u>underline</u> indicate the best and the second-best performing methods corresponding to the evaluation metrics. Rows are sorted in decreasing order of FScore (primary measure).

4. Methods for Task 1: Rhetorical Role Labeling

We received 26 runs from 11 teams for the task[1]. Table 1 compares the performance of the various runs. Brief descriptions of the methods are as follows. Details can be found in the working notes of the respective submissions.

- **rustic**[7]: The team from Huawei Ireland Research Centre, Dublin, Ireland was the best performing team in terms of F-Score and Recall. They treated the task as a Sequence Tagging problem, considering long-term label dependency between the sentences within a document. Along with that, structural, domain-specific, generic sentence embedding were used in different proportions. Bi-directional Gated Recurrent Unit (GRU) along with a Conditional Random Field (CRF) were used for tagging. They submitted 3 runs.

[1]This includes the teams that did not submit a working note

- **minitrue**: They used state-of-the-art transformer architectures such as LegalBert, RoBerta, and Bigbird along with a feed forward neural network for rhetorical role classification.
- **arguably**[8]: This team used ERNIE, ROBERTA as sequence classifier and experimented with pre-prossessing techniques like stop words removal, punctuations removal, lemmatizing and stemming.
- **ssn_nlp**[9]: This team too experimented with RoBERTa, LaBSE and BERT for the classification task.
- **nits_legal**[10]: They used Legal BERT along with MLP as a classifier. Moreover, oversampling strategy (SMOTE) for the minority class was followed.
- **legal_ai**: This team used recently published architecture of MiniLM to extract features, which were used to train multi-class SVM model.
- **ub_bw**[11]: This group experimented with fasttext classifier, its parameters and also with input type, namely, Unigram, Bigram and Trigram.
- **Chandigarh Concordia**: They attempted freezing and un-freezing all the layers of pretrained BERT and submitted two runs for that experiment. For the third run, they performed Sugeno Integral ensemble technique and the prediction prior runs as the input.
- **Legal_NLP**[12]: This team used cased and uncased BERT model experimented with the number of epochs, learning rate and other parameters.
- **cen-nlp**[13]: In this approach features are extracted from sentences using Distilroberta. They started experimenting with basic machine learning techniques and neural network architecture, and zero in on ANN as the best performing model. Furthermore, hyperparameter optimization performed by using GridSearchCV.
- **nit-agartala**[14]: This team used BERT and data pre-processing techniques for classification.

We find that the best performing method which achieved an FScore of 0.557 used BiDirectional GRU along with CRF. The best performing team also treated the task as Sequence Tagging same as current state of the art[4]. We observe that some variant of transformers or similar language models were widely used by almost all the teams, sometimes combined with other classifiers (SVM, FC etc.). Deep Learning methods that could extract deep semantic features were shown to perform much better than traditional feature based approaches. For several teams who also participated last year, the F1 score improved compared to the previous edition.

5. Methods for Task 2: Summarization of Legal Judgements

For the tasks of identifying 'summary-worthy' sentences in a court judgement (Task 2a), and automatically generating a summary from a given court judgement (Task 2b), we received a

Team Name	Run ID	P	R	F	Method Summary
Enigma	1	**0.64**	**0.58**	**0.59**	LegalBERT
nits_legal	2	0.61	0.57	<u>0.58</u>	LegalBERT and Multitask Objective
nits_legal	3	<u>0.64</u>	<u>0.58</u>	0.58	Ensemble using Run1 and Run2
nits_legal	1	0.63	0.57	0.57	MLP
NeuralMind	1	0.58	0.54	0.54	TextRank
NeuralMind	2	0.55	0.56	0.52	BM25
Chandigard_concordia	3	0.55	0.52	0.51	FAST AI TextRank on Term Document Frequency Matrix
Chandigard_concordia	2	0.55	0.56	0.50	BERT TextRank
NeuralMind	3	0.55	0.57	0.49	BM25
Chandigard_concordia	1	0.54	0.55	0.46	BERT TextRank on Cosine Similarity Matrix
nit_agartala_nlp_team	1	0.38	0.50	0.43	GCN

Table 2

Results of Task 2a: Summary-Worthiness Sentence Classification task. Measures averaged over 50 test documents. Numbers in **bold** and <u>underline</u> indicate the best and the second-best performing methods corresponding to the evaluation metrics. Rows are sorted in decreasing order of FScore (primary measure).

Team / Run ID	R1-P	R1-R	R1-F	R2-P	R2-R	R2-F	R4-P	R4-R	R4-F
nits_legal_1	0.647	<u>0.641</u>	**0.644**	<u>0.364</u>	<u>0.360</u>	**0.362**	<u>0.192</u>	**0.189**	**0.191**
nits_legal_3	0.643	0.637	<u>0.640</u>	0.360	0.357	<u>0.358</u>	0.190	<u>0.188</u>	<u>0.189</u>
nits_legal_2	0.642	0.635	0.639	0.355	0.351	0.353	0.183	0.181	0.182
NeuralMind_1	0.630	0.626	0.628	0.333	0.330	0.331	0.154	0.152	0.153
Chandigard_concordia_3	0.618	0.639	0.628	0.332	0.343	0.337	0.190	<u>0.188</u>	<u>0.189</u>
Chandigard_concordia_2	0.597	0.626	0.610	0.304	0.318	0.311	0.137	0.141	0.139
NeuralMind_3	0.601	0.595	0.592	0.299	0.299	0.297	0.130	0.131	0.130
nit_agartala_nlp_team_1	0.546	**0.668**	0.576	0.303	**0.370**	0.319	0.149	0.176	0.154
NeuralMind_2	0.574	0.578	0.564	0.282	0.284	0.278	0.119	0.120	0.118
Enigma_1	**0.680**	0.491	0.530	**0.394**	0.284	0.307	**0.208**	0.149	0.162
Chandigard_concordia_1	<u>0.658</u>	0.382	0.442	0.332	0.189	0.222	0.148	0.081	0.097

Table 3

Results of Task 2b: Legal Document Summarization in terms of ROUGE-Precision metric on test dataset. Measures averaged over 50 test documents. Numbers in **bold** and <u>underline</u> indicate the best and the second-best performing methods corresponding to the evaluation metrics. Rows are sorted in decreasing order of ROUGE-F1 Score (primary measure).

total of 11 runs from 5 participating teams[2]. The comparative results are in Table 2 (for Task 2a) and Table 3 (for Task 2b). We briefly describe below the methods used by each team in each of their runs. Details can be found in the working notes of the respective submissions.

- **enigma**[15]: This team treated the task as binary classification task, and used sentence embeddings from BERT pre-trained on legal corpus as features for Task 2a. For Task 2b, all the sentences that were classified as relevant in Task 2a were taken and concatenated together.

- **nits-legal**[10]: This team used legal pre-trained BERT and used sentence embeddings

[2]This includes the teams that did not submit a working note

generated from it as features for all the three runs submitted and also divided the training dataset in 5 shards. For Run 1, They trained different MLP models for these shards. Average of predictions from all model were considered. For Run 2, They employed the setup used in Run 1 to multitasking objective of Rhetorical Role labeling and Summary worthy sentence identification. In Run 3, They used of the models that are saved for Run1 and Run2. For Task 2b all the sentences that were classified as relevant were used to form a summary. They also experimented with different threshold to consider a sentence relevant.

- **neuralmind** : For the Task 2a, the team experimented with classical document ranking techniques, such TextRank and BM25. For Task 2b, Apart from classical techniques, they used recently published Long document summarizers, namely, LED and Pegasus. They applied these models on each document into segments of 1000 words and 500 words respectively and evaluated using zero-shot approach.

- **Chandigarh Concordia**: The team treated Task2a as binary classification task, and finetuned pre-trained Language Models using Fast AI and Tensorflow Libraries for submitting runs. For Task 2b, they relied on statistical approach of TextRank and applied to Term-Document Matrix, and Cosine-similarity Matrix, created from training data provided.

- **nit-agartala**[14]: This team used BERT pre-trained on legal corpus and GCN for Task 2a. For Task 2b, all the sentences that were classified as relevant in Task 2a were taken and concatenated together to form a summary.

For Task 2a (Identifying 'summary-worthy' sentences) the best performing run achieves a F1 Score of 0.59, where they used sentence embeddings from BERT pretrained on legal corpus as features. Approaches followed in Submitted Runs, can be broadly divided into classical approaches (e.g., TextRank, BM25) and Transformer based classifier or deep learning based summarizers (e.g., BERT pretrained on legal corpus, Pegasus, GCN etc.).

For Task 2(b) (Automatically generating a summary from a given court judgement) the best performing run achieves a ROUGE1-F1 Score of 0.644 (from nits-legal team), where they concatenate sentences which were relevant in Task 2a; this strategy was also followed by other teams such as nit_agartala_nlp_team and Enigma. Most of the teams filtered out relevant sentences using runs submitted in Task 2(a) and ordered them the way they arrive in the source legal document to form a summary. While extractive summaries are useful, the aim of a separate sub-task 2b was to encourage teams to attempt abstractive or generative summarization. However, most teams directly used the output of Task 2a as summary for Task 2b.

6. Conclusions

Like the previous two editions, AILA 2021 created new benchmark datasets and several systems for two important tasks related to legal data analytics. While we retained the task on rhetorical role labeling of sentences in Indian legal documents, a new legal document summarization task was introduced. For the rhetorical role labeling task we saw an improvement in performance compared to last year, which can in part be attributed to the task itself being a continuation and several teams from last year returning this year, and in part due to additional annotated data.

For the summarization task (Task 2), we expected more abstractive summaries, which was the key reason behind introducing a separate sub-task. However, most teams chose to concatenate the output of Task 2a and use it as a summary. This could be attributed, at least partially, to the small size of annotated corpus (500 documents), apart from the task itself being difficult.

Another observation this year is that most teams, across all tasks and subtasks, relied on a variant of pre-trained transformer models. This also enforces our belief that a larger training dataset will lead to a substantial improvement in the performance. We hope to offer these tasks again with a larger annotated dataset in future.

Acknowledgements: The track organizers thank all the participants for their interest in this track. We also thank the FIRE 2021 organizers for their support in organizing the track. The research is partially supported by TCG CREST through the project "Smart Legal Consultant: AI-based Legal Analytics".

References

[1] 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: Proceedings of FIRE 2019 - Forum for Information Retrieval Evaluation, 2019.

[2] 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.

[3] V. Parikh, U. Bhattacharya, P. Mehta, A. Bandyopadhyay, P. Bhattacharya, K. Ghosh, S. Ghosh, A. Pal, A. Bhattacharya, P. Majumder, AILA 2021: Shared task on artificial intelligence for legal assistance, in: D. Ganguly, S. Gangopadhyay, M. Mitra, P. Majumder (Eds.), FIRE 2021: Forum for Information Retrieval Evaluation, Virtual Event, India, December 13 - 17, 2021, ACM, 2021, pp. 12–15. URL: https://doi.org/10.1145/3503162.3506571. doi:10.1145/3503162.3506571.

[4] P. Bhattacharya, S. Paul, K. Ghosh, S. Ghosh, A. Wyner, Identification of rhetorical roles of sentences in indian legal judgments, in: Proc. International Conference on Legal Knowledge and Information Systems (JURIX), 2019.

[5] P. Bhattacharya, S. Poddar, K. Rudra, K. Ghosh, S. Ghosh, Incorporating domain knowledge for extractive summarization of legal case documents, in: Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL), 2021, p. 22–31.

[6] V. Parikh, V. Mathur, P. Mehta, N. Mittal, P. Majumder, Lawsum: A weakly supervised approach for indian legal document summarization, arXiv preprint arXiv:2110.01188v3 (2021).

[7] S. Dutta, Categorizing roles of legal texts via sequence tagging on domain-specific language models, in: FIRE 2021 (Working Notes), 2021.

[8] G. S. Kohli, P. Kaur, J. Bedi, Automatic detection of rhetorical role labels using ernie2.0 and roberta, in: FIRE 2021 (Working Notes), 2021.

[9] S. S. Balamurali, K. S, T. D, Simple transformers in rhetoric role labelling for legal judgements, in: FIRE 2021 (Working Notes), 2021.
[10] D. Jain, M. D. Borah, A. Biswas, Summarization of indian legal judgement documents via ensembling of contextual embedding based mlp models, in: FIRE 2021 (Working Notes), 2021.
[11] T. Leburu-Dingalo, E. Thuma, G. Mosweunyane, N. Motlogelwa, Rhetorical role labelling for legal judgements using fasttext classifier, in: FIRE 2021 (Working Notes), 2021.
[12] A. Mitra, Classification on sentence embeddings for legal assistance, in: FIRE 2021 (Working Notes), 2021.
[13] D. Sudharsan, A. U, P. B, S. K P, Distilroberta based sentence embedding for rhetorical role labelling of legal case documents, in: FIRE 2021 (Working Notes), 2021.
[14] S. Rusiya, A. Sharma, D. Debbarma, S. Debbarma, Rhetorical role labelling for legal judgements and legal document summarization, in: FIRE 2021 (Working Notes), 2021.
[15] S. Furniturewala, R. Jain, V. Kumari, Y. Sharma, Legal text classification and summarization using transformers and joint text features, in: FIRE 2021 (Working Notes), 2021.