Grenze International Journal of Engineering and Technology, Jan Issue

A Comprehensive Survey of Deep Learning Models for Legal Document Summarization

Kancharla Bharath Reddy[1] and Dr.J. Jayabharathy[2]
[1-2]Puducherry Technological University/Department of Computer Science and Engineering, Puducherry, India
Email: bharathreddy.k@ptuniv.edu.in, jayabharathy@ptuniv.edu.in

***Abstract*— Legal document Summarization poses a difficult challenge because of the complexity and diversity of legal language and concepts. In recent years, deep learning models have shown promising results in summarizing legal documents. The article examines the challenges of legal document summarization, reviews the state-of-the-art deep learning models, and identifies research gaps in this area. Potential future research directions and applications of legal document summarization using deep learning models are also discussed. The paper is a valuable resource for researchers and practitioners interested in legal document summarization using deep learning models.**

***Index Terms*— Legal Document, Text Summarization, Deep learning, Extractive, Abstractive Summarization.**

I. INTRODUCTION

A legal document is a written record that has legal significance and is enforceable by law. Legal document summarization aims to distil lengthy and complex legal texts into concise summaries, providing a quick understanding of the key information. This process is motivated by the need to improve efficiency in legal research, reduce information overload, and enhance access to legal knowledge. Summaries enable legal professionals to quickly identify relevant cases, arguments, and legal principles, saving time and effort. Additionally, they facilitate the communication of legal information to non-experts and aid in the automation of legal processes. Legal document summarization addresses the challenges of information overload and improves legal decision-making and analysis. The challenges of summarizing legal texts include capturing the nuances of legal language, condensing complex concepts, and ensuring accuracy and legal integrity. Effective legal document summarization improves productivity, facilitates legal research, and enhances communication of legal information.

Deep learning models have revolutionized natural language processing and text summarization by capturing semantic and contextual information from text. Models like recurrent neural networks (RNNs), transformers, and sequence-to-sequence architectures enable representation learning, sequence modelling, and encoder-decoder frameworks for generating summaries. Transformer models, such as BERT and GPT, leverage self-attention mechanisms to capture global dependencies and fine-tuning on large-scale datasets, while reinforcement learning techniques optimize summarization performance.

These models have been applied to both extractive and abstractive summarization, benefiting from pre-training on massive corpora and transfer learning. Overall, deep learning models have significantly advanced NLP and summarization tasks by improving summary quality, language understanding, and information extraction.

Grenze ID: 01.GIJET.10.1.305

A. Deep Learning Approaches and Methodologies in Legal Document Summarization

Deep learning concepts and architectures relevant to legal document summarization include recurrent neural networks (RNNs), convolution neural networks (CNNs), transformers, and their variants. RNNs are suited for sequential data processing, while CNNs excel in feature extraction and image processing. Transformers are particularly useful for capturing long-range dependencies and generating abstractive summaries. Other important concepts include attention mechanisms, transfer learning, and reinforcement learning. These models can be adapted and modified to handle legal texts and have shown great potential in improving the accuracy and efficiency of legal document summarization.

Adapting deep learning models for legal texts involves incorporating legal domain knowledge through pre-trained legal word embeddings or creating legal ontologies. For example, a model trained on legal contracts can learn to understand specific legal terms and clauses. Fine-tuning the models on legal-specific datasets helps them grasp the context and nuances of legal jargon, such as "breach of contract" or "reasonable doubt." Attention mechanisms enable the models to focus on key sections within legal documents, while contextual embeddings like BERT capture the hierarchical structure and relationships between different clauses. Additionally, privacy and ethical concerns are addressed through techniques such as differential privacy and anonymization to protect sensitive legal data.

Recurrent Neural Networks (RNNs) excel in sequential data processing by capturing dependencies over time, making them suitable for modelling the sequential nature of legal texts. Convolution Neural Networks (CNNs) are adept at extracting local features and patterns from text, enabling them to identify important textual elements in legal documents. Transformers, with their attention mechanisms, capture global dependencies and have revolutionized natural language processing, including legal text summarization. Variants of these models, such as Bidirectional RNNs, Dilated CNNs, and Transformer-based architectures like BERT, have further enhanced their capabilities in understanding and summarizing legal texts. Each model and its variants offer unique strengths in processing legal documents and generating accurate summaries.

In the field of legal document summarization using deep learning, training data plays a crucial role in the development and evaluation of models. Annotated datasets specifically created for legal texts are essential for training deep learning models to generate accurate summaries. However, the creation of such datasets presents challenges, including the need for expert annotators with legal domain expertise and the careful consideration of legal privacy concerns. Furthermore, evaluating the quality of generated summaries requires appropriate evaluation metrics. Commonly used metrics in the field include ROUGE (Recall-Oriented Understudy for Gisting Evaluation) and BLEU (Bilingual Evaluation Understudy), which assess the overlap and similarity between generated summaries and reference summaries.

B. Techniques for Extractive and Abstractive Summarization

Techniques for Extractive and Abstractive Summarization play a significant role in legal document summarization. Extractive methods involve identifying and selecting important sentences or segments from the source document to form the summary. This can be achieved through sentence scoring techniques, where sentences are ranked based on relevance or importance scores. Named entity recognition is employed to identify and include important entities such as key people, organizations, or legal concepts. Keyword extraction techniques identify crucial keywords and include them in the summary to provide context and relevance.

On the other hand, abstractive methods aim to generate summaries by understanding the source document's meaning and generating new, concise sentences that capture the essential information. Sequence-to-sequence models, such as encoder-decoder architectures, have been successfully adapted to abstractive summarization. These models learn to encode the source document into a representation and generate a summary based on that representation. Attention mechanisms, often used in conjunction with sequence-to-sequence models, help the model focus on relevant parts of the source document during summary generation.

Abstractive methods, on the other hand, have the advantage of generating more concise and coherent summaries by paraphrasing and combining information from multiple parts of the document. They can capture the essence of the source document and provide more human-like summaries. However, abstractive methods can be more challenging to train and may sometimes introduce errors or produce summaries that deviate from the original document's context.

II. Related Works

Anastassia Kornilova et. Al [1] introduced the BillSum dataset, which is the first corpus for legislative summarization. The dataset contains US Congressional bills and California state bills, split into a train and a test

set. The paper establishes several benchmarks for automatic legislative summarization and shows that there is ample room for new methods that are better suited to summarize technical legislative language. The paper also demonstrates that summarization methods trained on US bills transfer to California bills, indicating that the methods developed on this dataset could be used for legislatures without human-written summaries.

Linwu Zhong et.al [2] used a novel train-attribute-mask pipeline and a sentence type classifier to select predictive sentences from legal case documents, and then uses Maximum Marginal Relevance to select summarization sentences. The generated summaries are evaluated using ROUGE metrics and a qualitative survey. The paper also highlights the challenges of summarizing legal cases and suggests that future work should focus on case-aspect coverage and explore how large collections of cases enable a data-driven inference about which parts of the text are relevant for the case outcome to reduce the need for domain-specific model building.

Vedant Parikh et. al [3] proposed a new dataset of Indian legal documents and their summaries, which can be used for legal analytics research. The automatic labelling technique for identifying summary-worthy sentences and the weakly supervised sentence extractor achieve high accuracy in summarization.

Paheli Bhattacharya et. al [4] Introduced an unsupervised summarization algorithm named DELSumm, which systematically integrates domain expertise for the purpose of extracting summaries from legal case documents. The suggested algorithm demonstrates superior performance in ROUGE scores compared to various robust benchmark methods, encompassing both general summarization techniques and those tailored for legal contexts.

Praveen Bushipaka et. al [7] presented a legal holding extraction method based on Italian-LEGAL-BERT to automatically extract legal holdings from Italian cases. The authors also introduce ITA-CaseHold, a benchmark dataset for Italian legal summarization. They set new baselines for this dataset by introducing HM-BERT, a new extractive summarization tool based on Italian-LEGAL-BERT. The experiments conducted using this dataset show that the harmonic mean improves the effectiveness of the model with respect to the arithmetic mean.

Marios Koniaris et. al [8] evaluated automatic summarization of Greek legal documents using several state-of-the-art methods for abstractive and extractive summarization. The paper presents a new metadata-rich dataset consisting of selected judgments from the Supreme Civil and Criminal Court of Greece, alongside their reference summaries and category tags, tailored for the purpose of automated legal document summarization. The paper also identifies the need for better metrics that can capture the coherence, relevance, and consistency of legal document summaries.

Deepali Jain et. al [9] proposed a novel sentence scoring approach called DCESumm for legal document summarization. It involves supervised sentence-level summary relevance prediction and unsupervised clustering-based document-level score enhancement. The approach uses Legal BERT and deep clustering techniques to significantly improve the scores of summary-relevant sentences. Experimental results on two legal document summarization datasets, BillSum and Forum of Information Retrieval Evaluation (FIRE), show that the proposed approach outperforms the current state-of-the-art approaches in terms of ROUGE metric F1-score improvements.

Aniket Deroy et. al [10] examined multiple methods of combining models and demonstrated that several of their ensemble approaches generate summaries that surpass those produced by individual foundational algorithms, as measured by ROUGE and METEOR scores. They also introduce innovative graph-based ensemble techniques that utilize graph centrality metrics, achieving superior results across all base algorithms in numerous scenarios.

Paheli Bhattacharya et. al [11] presented a study that compares various summarization algorithms when applied to legal case rulings is conducted by the authors. They conduct experiments using an extensive collection of Indian Supreme Court judgments and a diverse range of summarization algorithms, including unsupervised and supervised methods. The objective is to evaluate the effectiveness of domain-independent summarization techniques on legal case judgments and to determine how approaches tailored for legal documents from other countries (like Canada and Australia) can be applied to Indian Supreme Court materials. In addition to quantitative assessment by comparing with reference summaries, the study also provides valuable qualitative observations on algorithm performance from the viewpoint of a legal expert

Deepa Anand et. al [21] proposed basic and versatile neural network methods for summarizing Indian legal case documents. The suggested methods operate independently of manually designed attributes or specialized domain expertise, and their applicability is not confined to a specific sub-field, rendering them adaptable to diverse domains. Addressing the scarcity of labeled data, the issue is resolved by assigning classes/scores to sentences in the training dataset, gauged against the alignment with human-generated reference summaries. Through experimental assessments, the effectiveness of the proposed techniques is established in comparison to alternative foundational methods.

The survey on Legal document summarization using deep learning models has certain limitations that need to be addressed in the future. Firstly, deep learning models require a large amount of annotated training data, which can

be scarce in the legal domain due to limited access to labeled legal documents. To overcome this, efforts should be made to curate larger and more diverse annotated datasets specific to legal texts. Secondly, deep learning models often struggle with capturing legal nuances, such as legal jargon, context, and precedent. This can lead to inaccuracies in summarization. To address this, future models could incorporate legal domain knowledge and leverage techniques like pre-training on legal corpora or fine-tuning on legal-specific tasks. Additionally, the interpretability of deep learning models remains a challenge in the legal domain. Developing methods to provide explanations for model-generated summaries could enhance transparency and trustworthiness. Lastly, ethical concerns related to privacy, data security, and bias need to be carefully addressed to ensure the responsible and fair deployment of these models in legal settings.

S.no	Author	Title	Techniques	Dataset	Metrics	Limitations
1	Anastassia Kornilova[1]	BillSum: A Corpus for Automatic Summarization of US Legislation, arXiv:1910.0052,2019	Neural sentence representations and traditional contextual features	BillSum	ROUGE-1, ROUGE-2, and ROUGE-L	Limited to bills from US Congress and California state. Only explores extractive summarization methods.
2	Linwu Zhong[2]	ICAIL '19Automatic Summarization of Legal Decisions using Iterative Masking of Predictive Sentences, ICAIL '2019	CNN classifier	Post-traumatic Stress Disorder (PTSD)	ROUGE-1, ROUGE-2	Summary does not always cover all decision-relevant aspects of a case.
3	Varun Pandya	Automatic Text Summarization of Legal Cases: A Hybrid Approach, arXiv:1908.09119, 2019	k-means clustering technique and tf-idf word vectorizer	Auslii	ROGUE 1, ROGUE 2, ROGUE L, ROGUE W	Does not consider the semantic meaning of sentences. May not work well for cases with a large number of documents.
4	Vedant Parikh[3]	LawSum: A weakly supervised approach for Indian Legal Document Summarization, arXiv:2110.01188, 2021	Bidirectional-LSTM	Supreme Court of India Judgements	ROGUE	Automatic labeling technique may not capture all important information.
5	Paheli Bhattacharya[4]	Incorporating Domain Knowledge for Extractive Summarization of Legal Case Documents, ICAIL '21	DELSumm	50 Indian Supreme Court case documents	ROUGE	Relatively small dataset of 50 Indian Supreme Court case documents may limit generalizability to other legal domains or languages.
6	Yuxin Huang[5]	Exploiting comments information to improve legal public opinion news abstractive summarization, Front. Comput. Sci., 2022	Hierarchical comment-aware encoder (HCAE)	Article-summary-comments pairs from MicroBlog sites	ROUGE-1, ROUGE-2, and ROUGE-L	Lack of detailed analysis of the impact of the selective denoising module on model performance.
7	Abhay Shukla[6]	Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation, arXiv:2210.07544, 2022	Reduction, LexRank, LSA, DSDR, PacSum, SummaRuN-Ner and BERTSum	Indian-Abstractive (IN-Abs), Indian-Extractive (IN-Ext), US-Extractive (US-Ext)	ROUGE-1, ROUGE-2, ROUGE-L, BERTScore	Transformer-based abstractive summarization models have restrictions on input tokens. Automatic summarization metrics may not fully capture quality in

						specialized domains like law.
8	Praveen Bushipaka[7]	Legal Holding Extraction from Italian Case Documents using Italian-LEGAL-BERT Text Summarization, ICAIL-2023	Italian-BERT, Italian-LEGAL-BERT, Italian-LEGAL-BERT-SC models	ITA-CaseHold	ROUGE-1, ROUGE-2, ROUGE-L, ROUGE-W	Model may select redundant information due to syntactic overlap with corresponding document holding.
9	Marios Koniaris[8]	Evaluation of Automatic Legal Text Summarization Techniques for Greek Case Law, Information, 2023 - mdpi.com	LetSum, CaseSummarizer, graph-based, BERT-based models	Judgments from the Supreme Civil and Criminal Court of Greece	ROUGE-1, ROUGE-2, ROUGE-3, ROUGE-L, ROUGE-W	Limited evaluation of summarization methods.
10	Deepali Jain[9]	A sentence is known by the company it keeps: Improving Legal Document Summarization Using Deep Clustering, Artificial Intelligence and Law, 2023	Legal BERT embeddings, deep embedded clustering-based approach (DEC)	BillSum, Forum of Information Retrieval Evaluation (FIRE)	ROUGE-S, ROUGE-W, ROUGE-L, ROUGE-N	May not always be feasible in real-world scenarios. Effectiveness may be limited in documents from other domains.
11	Aniket Deroy[10]	Ensemble methods for improving extractive summarization of legal case judgements, Artif Intell Law 2023	Ensembling techniques	Court case judgements from the Indian Supreme Court	ROUGE and METEOR scores	Effectiveness of ensemble techniques may depend on quality and diversity of base summarization algorithms used.
12	Paheli Bhattacharya[11]	A Comparative Study of Summarization Algorithms Applied to Legal Case Judgments, Advances in Information Retrieval, 2019	Luhn's method, graph-based approaches, matrix-based approaches, data reconstruction approach (DSDR)	Legal case documents from the Supreme Court of India	ROUGE, F1 score, precision, recall, sentence-level coherence	Indian case documents lack fixed structure and section headings. Citations to statutes are important for summaries of Indian legal case judgments.
13	Deepali Jain[12]	Bayesian Optimization based Score Fusion of Linguistic Approaches for Improving Legal Document Summarization, Knowledge-Based Systems, 2023	Bayesian Optimization (BO) based Score Fusion	BillSum, Intelligence for Legal Assistance (AILA)	ROUGE-1, ROUGE-2, ROUGE-L	Proposed approach may not always be feasible or practical in real-world scenarios.
14	Diego de Vargas Feijo et al[13]	Improving abstractive summarization of legal rulings through textual entailment, LegalSumm	LegalSumm	RulingBR	ROUGE	Summaries may be limited in length if needed to be twice as large.

III. Conclusion

In conclusion, this survey explored the effectiveness of deep learning models in legal document summarization. These models have demonstrated promising results in capturing the nuances of legal language and context. However, the survey also identified several research gaps and limitations that need to be addressed. These include the lack of explainability in deep learning models, the scarcity of labeled training data, difficulties in handling complex legal language, and challenges in summarizing long documents. Overcoming these limitations will contribute to the development of more accurate, reliable, and interpretable legal document summarization systems.

Future research directions should focus on addressing the identified limitations and developing hybrid approaches that combine deep learning models with other techniques such as rule-based systems or knowledge graphs to enhance the summarization process. Additionally, the development of domain-specific evaluation metrics and frameworks will provide more robust measures to assess the quality of legal document summaries.

REFERENCES

[1] Kornilova, Anastassia, and Vlad Eidelman. "BillSum: A corpus for automatic summarization of US legislation." arXiv preprint arXiv:1910.00523 (2019).

[2] Zhong, Linwu, Ziyi Zhong, Zinian Zhao, Siyuan Wang, Kevin D. Ashley, and Matthias Grabmair. "Automatic summarization of legal decisions using iterative masking of predictive sentences." In Proceedings of the seventeenth international conference on artificial intelligence and law, pp. 163-172. 2019.

[3] Parikh, Vedant, Vidit Mathur, Parth Mehta, Namita Mittal, and Prasenjit Majumder. "Lawsum: A weakly supervised approach for indian legal document summarization." arXiv preprint arXiv:2110.01188 (2021).

[4] Bhattacharya, Paheli, Soham Poddar, Koustav Rudra, Kripabandhu Ghosh, and Saptarshi Ghosh. "Incorporating domain knowledge for extractive summarization of legal case documents." In Proceedings of the eighteenth international conference on artificial intelligence and law, pp. 22-31. 2021.

[5] Huang, Yuxin, Zhengtao Yu, Yan Xiang, Zhiqiang Yu, and Junjun Guo. "Exploiting comments information to improve legal public opinion news abstractive summarization." Frontiers of Computer Science 16, no. 6 (2022): 166333.

[6] Shukla, Abhay, Paheli Bhattacharya, Soham Poddar, Rajdeep Mukherjee, Kripabandhu Ghosh, Pawan Goyal, and Saptarshi Ghosh. "Legal case document summarization: Extractive and abstractive methods and their evaluation." arXiv preprint arXiv:2210.07544 (2022).

[7] Licari, Daniele, Praveen Bushipaka, Gabriele Marino, Giovanni Comandé, and Tommaso Cucinotta. "Legal Holding Extraction from Italian Case Documents using Italian-LEGAL-BERT Text Summarization." (2023).

[8] Koniaris, Marios, Dimitris Galanis, Eugenia Giannini, and Panayiotis Tsanakas. "Evaluation of Automatic Legal Text Summarization Techniques for Greek Case Law." Information 14, no. 4 (2023): 250.

[9] Jain, Deepali, Malaya Dutta Borah, and Anupam Biswas. "A sentence is known by the company it keeps: Improving Legal Document Summarization Using Deep Clustering." Artificial Intelligence and Law (2023): 1-36.

[10] Deroy, Aniket, Kripabandhu Ghosh, and Saptarshi Ghosh. "Ensemble methods for improving extractive summarization of legal case judgements." Artificial Intelligence and Law (2023): 1-59.

[11] Bhattacharya, Paheli, Kaustubh Hiware, Subham Rajgaria, Nilay Pochhi, Kripabandhu Ghosh, and Saptarshi Ghosh. "A comparative study of summarization algorithms applied to legal case judgments." In Advances in Information Retrieval: 41st European Conference on IR Research, ECIR 2019, Cologne, Germany, April 14–18, 2019, Proceedings, Part I 41, pp. 413-428. Springer International Publishing, 2019.

[12] Jain, Deepali, Malaya Dutta Borah, and Anupam Biswas. "Bayesian Optimization based Score Fusion of Linguistic Approaches for Improving Legal Document Summarization." Knowledge-Based Systems 264 (2023): 110336.

[13] Feijo, Diego de Vargas, and Viviane P. Moreira. "Improving abstractive summarization of legal rulings through textual entailment." Artificial intelligence and law 31, no. 1 (2023): 91-113.

[14] Shukla, Bharti, et al. "Text summarization of legal documents using reinforcement learning: A study." Intelligent Sustainable Systems: Proceedings of ICISS 2022. Singapore: Springer Nature Singapore, 2022. 403-414.

[15] Satwick Gupta, M. N., et al. "Extractive summarization of Indian legal documents." Edge Analytics: Select Proceedings of 26th International Conference—ADCOM 2020. Singapore: Springer Singapore, 2022.

[16] Klaus, Svea, et al. "Summarizing legal regulatory documents using transformers." Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022.

[17] Aumiller, Dennis, Ashish Chouhan, and Michael Gertz. "EUR-lex-sum: A multi-and cross-lingual dataset for long-form summarization in the legal domain." arXiv preprint arXiv:2210.13448 (2022).

[18] Jain, Deepali, Malaya Dutta Borah, and Anupam Biswas. "Summarization of legal documents: Where are we now and the way forward." Computer Science Review 40 (2021): 100388.

[19] Kanapala, Ambedkar, Sukomal Pal, and Rajendra Pamula. "Text summarization from legal documents: a survey." Artificial Intelligence Review 51 (2019): 371-402.

[20] Rusiya, Siddhartha, and Anupam Jamatia. "Implementation of Legal Documents Text Summarization and Classification by Applying Neural Network Techniques." Machine Intelligence Techniques for Data Analysis and Signal Processing: Proceedings of the 4th International Conference MISP 2022, Volume 1. Singapore: Springer Nature Singapore, 2023.

[21] Deepa Anand, & Rupali Wagh. (2019). Effective deep learning approaches for summarization of legal texts. Journal of King Saud University - Computer and Information Sciences, ISSN 1319-1578.

[22] Jain, D., Borah, M. D., & Biswas, A. (2021). Summarization of legal documents: Where are we now and the way forward. Computer Science Review, 40, 100388.