LegalSummNet: A Transformer-Based Model for Effective Legal Case Summarization

Md Farhad Kabir[1], Sohana Afrin Mitu[2], Sharmin Sultana[3], Belal Hossain[4], Rakibul Islam[5], Khandakar Rabbi Ahmed[6]*
Marshall School of Business, University of Southern California, Los Angeles, United States [1]
Department of Management, International American University, Los Angeles, United States [2]
Department of Business Administration and Management, International American University, Los Angeles, United States [3]
Department of Management Information System, International American University, Los Angeles, United States [4]
Department of MBA in Business Analytics, International American University, Los Angeles, United States[5]
Adjunct Researcher, Miyan Research Institute, International University of Business Agriculture and Technology, Bangladesh [6]

***Abstract*—Expanding legal documents has become increasingly complicated and presents a greater challenge to legal professionals in extracting relevant information efficiently. In this paper, a new two-stage hybrid summarization system, called LegalSummNet, is introduced. It excels in handling the peculiarities of legal texts, such as their extremely long length, complex syntax, and specialized vocabulary. LegalSummNet combines an extractive filtering model with an attention-weighted filtering module and a transformer-based abstractive generation model, enabling it to identify significant elements and produce compact, coherent, and semantically competent summaries. The proposed model is tested using a large-scale dataset comprising a legal case and shows significant improvements compared to robust baselines, such as BERTSumExt and LegalT5, in performance measured by ROUGE-1, ROUGE-2, ROUGE-L, and BERT Score. A greater compression efficiency is also evident with the model. Hence, there is a significant application of real-world systems in generating case briefs and summaries related to contracts. The findings demonstrate that LegalSummNet is effective in enhancing the accessibility of legal documents and supporting informed decision-making.**

***Keywords*—Legal document summarization; NLP, extractive and abstractive summarization; transformer; LegalSummNet; BERT; LegalT5; ROUGE-L**

I. INTRODUCTION

The practice of law is increasingly affected by the rapid proliferation of computerized legal materials, including statutes, case law, court rulings, and administrative regulations. Legal practitioners—judges, attorneys, clerks, and scholars—must contend with the growing volume of information to extract legally pertinent content for research, case preparation, and judicial rulings. In practice-based areas such as commercial litigation and multi-party contract cases, legal professionals must navigate numerous pages of filings, judgments, and supporting documents that must be interpreted within tight timelines [1].

This information overload, combined with the diffuse structure and technical language of legal writings, makes it far more likely to be misinterpreted and wasteful in legal processes. Legal case summaries provide a solution by condensing lengthy documents into concise, easy-to-understand representations of case facts, arguments, and decisions. These summaries reduce cognitive loads, expedite legal review, and promote consistency in legal research and judgment. However, producing them is still accomplished mainly by hand—an operation that is time-consuming, error-prone, and unworkable in high-volume situations. Automated summarization, therefore, appears to be a crucial tool for enhancing the accessibility, efficiency, and accuracy of legal analysis. It is, however, a daunting task to develop such tools for the law profession. Legal texts not only consist of lengthy and lexically dense language, but they must also meet strict formal requirements in terms of legal terminology, interpretability, and logical composition [2].

Traditional extractive summarization methods, such as Latent Semantic Analysis (LSA)-based approaches, have shown promise in reducing document length and isolating key sentences. However, they are insensitive to deep, implicit semantic relationships and fail to retain the argumentative organization or factual accuracy of legal narratives. More recently, pre-trained language models such as BERT, T5, and PEGASUS have transformed natural language processing (NLP) [3] by enabling context-sensitive language understanding and generative summarization. Legal-specific models, such as LegalBERT and LegalT5, have further built upon these strengths, delivering improved performance when applied to legal documents. Domain-specific data set availability—most notably the over 7,900 document-summary pair Legal Case Document Summarization dataset—has enabled the creation of data-driven models for legal summarization [4], [5].

Despite these innovations, abstractive models are usually deficient in two areas: (1) token-length input limitations that make long legal documents processable as a whole, and (2) generating summaries that are grammatically perfect but legally Insufficient or incorrect. Furthermore, standard evaluation metrics such as ROUGE or BLEU are also found to be unsuitable for measuring the legal fidelity and usability of synthesized summaries. These gaps indicate the need for a hybrid approach that not only selects and screens legally pertinent information but also generates coherent, concise, and domain-suitable summaries [6].

To solve these problems, we propose LegalSummNet, a novel hybrid system for legal case text summarization. LegalSummNet is inspired by how legal professionals natively summarize: initially extracting pertinent content (extractive filtering) and subsequently paraphrasing the content in condensed form (abstractive generation). Our system integrates BERTSumExt for sentence-level scoring and filtering with LegalT5, a transformer fine-tuned specifically to the legal task, for abstractive summary generation. This two-stage pipeline overcomes token-length issues, reduces noise, and draws more attention to legally salient content.

Our key contributions to this research are as follows:

- We introduce LegalSummNet, a hybrid model that mimics human-like summarization behavior by combining extractive and abstractive stages, tailored explicitly for legal texts.
- We implement an attention-weighted scoring mechanism using BERTSumExt to select the most relevant sentences from long legal documents. This step effectively reduces input size while preserving semantic integrity.
- We introduced an extractive filtering stage, which reduces the input sequence length L relatively before feeding it into the Transformer-based abstractive model.

The rest of the paper is organized as follows: Section II is a comprehensive review of current research in legal document summarization and evaluation metrics. Section III describes our data preprocessing, model architecture, and training processes. Section IV presents experimental results, including quantitative metrics and qualitative findings. Section V concludes with discussions of the study's implications, limitations, and future work.

II. LITERATURE REVIEW

This section examines the evolution of computerized legal text summarization techniques, with a particular focus on the significant contributions from traditional machine learning methods to advance transformer-based and hybrid explainable AI methods. Particular attention is given to how these works meet the idiosyncrasies of legal texts — their complexity, length, jargonized vocabulary, and interpretability demands. Through questioning them, this section maps advances made, challenges encountered, and potential avenues of future advancement in automatic summarization and legal document analysis. During the early stages of this research line, heuristic-based methods and conventional machine learning were the pillars of legal summarization systems.

Alpana et al. [7] proposed a machine learning technique to identify and summarize legal arguments from legal documents with an accuracy rate of 0.94. Their approach dramatically reduced the workload of legal professionals by alleviating time-consuming summarization tasks, demonstrating that domain adaptation enhances automatic processing.

Furthermore, Kesavarapu et al. [8] also developed a web application that utilized the Text Rank unsupervised algorithm to generate extractive summaries from lengthy case files. Their application automated judicial proceedings by presenting summaries of key case information, thereby reducing the cognitive load on decision-makers.

Bhattacharya et al. [9] further developed this concept in the form of DEL Sum, an unsupervised summary system specifically designed to incorporate expert legal knowledge into the summarization process. It surpassed many supervised baselines for Indian Supreme Court judgments, demonstrating the effectiveness of domain knowledge embedding in enhancing the quality and relevance of a summary.

The employment of latent semantic analysis (LSA) by Merchant and Pande [10] was a sterling example of the capability of semantic concept extraction in generating short, understandable summaries to a level of professional acceptability, as evidenced by high ROUGE-1 test scores. These path-breaking research papers collectively highlight that machine learning, combined with domain expertise, is a key factor in bridging the complexity gap in legal documents. A paradigm shift was brought about with the coming of deep learning. Transformer models, such as BERT, Longformer, BART, and PEGASUS, have changed the face of summarizing legal texts.

The efficiency of fine-tuning the Longformer Encoder-Decoder (LED) model on data from Pakistani and Australian judiciaries was demonstrated by Sarwar et al. [11] with a very high ROUGE-1 score of more than 0.53. This result demonstrates the ability of long-formers to work with long texts with successful modeling of long-distance relationships—a feature of much significance for legal texts, where coherence in context is of such overriding concern.

Kasar et al. [12] had compared BART and PEGASUS models trained on US Congressional bills as a legal text proxy. They preferred BART, which performed better in identifying legal content structures and producing coherent summaries. Barros et al. [13] utilized BERT's contextual embeddings for the extractive summarization of Brazilian Federal Police reports, demonstrating the model's capacity to process complex, domain-specific investigative reports.

Prabhakar and Pati [14] compared the performances of both T5 and GPT-2 generative transformer models in Indian court verdicts, with T5 outperforming GPT-2 substantially based on ROUGE and BERT Score. Both of these papers serve as a gold standard for the pioneering efforts of pre-trained transformers, which can comprehend intricate legal terminology and produce high-quality summaries with minimal human feature engineering.

Despite their success, transformer models have some inherent limitations when used in legal text summarization tasks, such as token length constraints and the black-box properties of deep models. To resolve these issues, Turan and Küçüksille [15] developed a new hybrid summarization tool that combines extractive and abstractive approaches, particularly for Turkish Constitutional Court rulings. Not only did they remove token length constraints, but they also improved summary quality, achieving ROUGE scores of nearly 0.6. Aside from summarization, they constructed XGBoost-powered classifiers to forecast legal decisions, which were explained using SHAP explainability to inform model choices—a significant step toward establishing trust and transparency in legal AI systems.

Yao, Bai, and Liu [16] also developed a combined network model for multi-oriented text detection and recognition. Although designed for natural scenes, it has significant relevance in the digitization of legal documents. Their solution addresses mixed orientation and font issues in scanned legal documents, enabling the accurate extraction of text, a basic necessity for effective summarization. This innovation aligns with the growing consensus that hybrid and explainable AI methodologies, as well as domain adaptation, are essential for creating stable, interpretable, and beneficial legal NLP tools.

In recent works, summarization approaches have been devised to lower the dimensionality of feature space, which subsequently enhances the performance of text genre classification. Basha et al. proposed a technique for multi-document summarization with sentence importance, which takes into consideration the term frequency and sentence similarity for better feature selection and classification accuracy [20].

In short, there is a visible shift in the literature from heuristic, domain-specific summarization techniques to sophisticated deep models grounded in hybrid support and interpretability frameworks. The earlier methods suggested that summarization models would need to incorporate legal knowledge to mitigate the domain's complexity. Subsequent transformer-based methods have significantly improved performance by utilizing pre-trained language models to capture fine-grained semantic and syntactic patterns latent in legal texts. Still, it is challenging to support long and heterogeneous legal documents, ensure transparency, and deploy them in real-world legal applications.

Therefore, future work should focus on designing hybrid architectures that robustly combine domain expertise with powerful transformer models, while also incorporating explainability methods, to bridge the gap between AI performance and the demands of trust and usability from lawyers.

III. Methodology

This section outlines the entire methodology of this research, which proposes a solution based on a two-stage hybrid architecture designed to process the extreme length and complexity of legal documents. The entire process, as illustrated in Fig. 1, begins with the preprocessing of raw legal case data, followed by an extractive content filtering phase and an abstractive summarization phase. It firstly employs an attention-weighted mechanism to detect salient sentences, and finally, the coherent summaries are generated based on a transformer-based encoder-decoder. The final results are evaluated using typical summarization metrics and analyzed for efficiency. All the elements of this work process are explained in detail below.

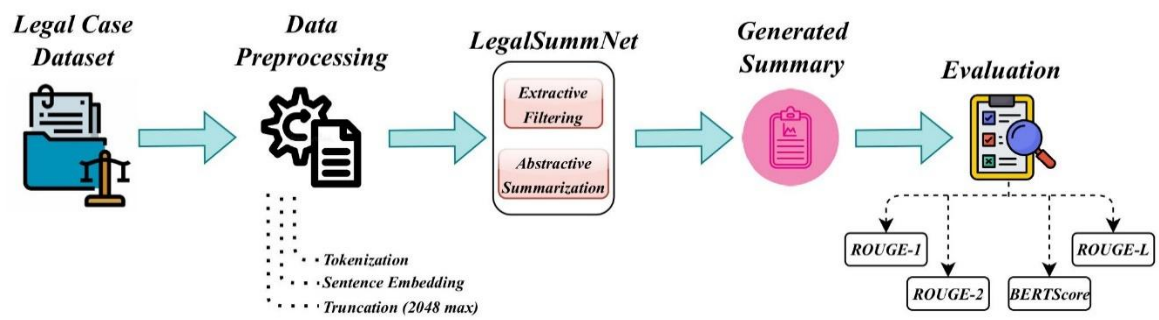

Fig. 1. Workflow of the research.

A. Dataset Description

The Legal Case Document Summarization dataset [17] is utilized in this study, which was collected from Kaggle and contains 7,973 pairs of documents and their corresponding summaries, all of which were composed by a person. All the case documents and summaries are in the form of .txt, which is the proper format to use in extractive and abstractive summary tasks.

In this rich set of practical legal cases, the given resource will provide a valuable training and test dataset for an NLP model, enabling it to produce concise and precise summaries of legal texts. Legal practitioners can utilize these summaries to navigate the vast volumes of legal literature.

B. Dataset Preprocessing

The dataset consists of .txt files that include full legal documents $X = \{x^{(1)}, x^{(2)}, ..., x^{(N)}\}$ and human written summaries $Y = \{y^{(1),} y^{(2)}, ..., y^{(N)}\}$, where $x^{(i)} \in R^{L}$ is legal text and $y^{(i)} \in R^{M}$ is natural text, and typically L" M.

1) Sentence segmentation and tokenization: Each of the documents $x^{(i)}$ gets divided into a set of sentences:

$$x^{(i)} = \{s_1^{(i)}, s_2^{(i)} \, ...s_n^{(i)} \} \tag{1}$$

Each sentence s_j is tokenized by a domain-specific tokenizer to interpret legal terms, as:

$$s_j = \{w_1^{(j)}, w_2^{(j)} \, , ...s_{T_j}^{(j)} \} \tag{2}$$

2) Named entity recognition and text cleaning: Legal-specific NER is carried out to extract entities $E = \{e_1, e_2, ..., e_k\} \in \varepsilon$, where ε is the set of all legal entities (judges, statutes, case citations, etc.). Normalization of documents is achieved through the elimination of boilerplate text and lemmatization, without excluding legal stopwords.

3) Embedding representation: Legal-BERT encodes each sentence sj into an embedding:

$$h_j = BERT(s_j) \in \mathbb{R}^d, d = 768 \tag{3}$$

The document is then encoded as a sequence of contextual embeddings:

$$H = [h_1, h_1, \ldots h_1,] \in \mathbb{R}^{n*d} \quad (4)$$

C. Baseline Model

1) BERTSumExt: Extractive summarization is addressed as a binary sequence labeling. The target is to assign a binary label $y_j \in \{0, 1\}$ that can be used to indicate whether sentence s_j is included in the summary, given a set of sentence embeddings H. BERTSumExt adds on BERT a classification layer predicting the importance of a sentence. Supposing the embedding of sentence s_j is denoted by h_j:

$$p_j = \sigma\{w_c . h_j + b_c), \ p_j \epsilon [0,1] \quad (5)$$

where the sigmoid activation is denoted by $\sigma(\cdot)$, and the trainable parameters are by $W_c \in R^{1\times d}$, and $b_c \in R$.

To maximize the extractive summarization model, the training objective is defined as a binary classification problem, where the model determines whether to include or exclude a sentence in the summary. The binary cross-entropy is used to train the model, which can be expressed as:

$$\mathcal{L}_{ext} = -\sum_{j=1}^{n} [y_j \log(p_j) + (1 - y_j) \log(1 - p_j)] \quad (6)$$

Where $y_j \in \{0, 1\}$ is the ground truth label of sentence sj and $p_j \in [0, 1]$ is the probability of inclusion. The loss penalizes the model more in cases of incorrect prediction, especially when the model assigns high probabilities to incorrect labels.

2) LegalT5: The task of abstractive summarization is represented as a sequence-to-sequence mapping, $x \rightarrow y$, so that the model can produce a coherent summary by learning semantics of language and discourse structure. LegalT5 is a domain-specific variation of Text-to-Text Transfer Transformer (T5). It consists of:

- Encoder: The encoder takes as input tokens x = {x1, . . . , xL} and maps them to hidden states Z = {z1, . . ., zL}, in which every hidden state can be calculated as:

$$z_i = LayerNorm(SelfAttn(x_i) + x_i \quad (7)$$

where Self Attention $(\cdot)$ represents the multi-headed self-attention mechanism that captures dependencies among tokens of the input, and the residual connection $+x_i$ helps to retain the original features of the input. Layer Norm is a normalization operation applied to stabilize training.

- Decoder: Output summary token $\hat{y}_t$. At any timestep t, the decoder generates the token autoregressively based on previous tokens generated up to and including y<t and the encoder output Z.

$$\hat{y}_t = \arg\max_{y \in V} P(y_t | y_{<t}, Z) \quad (8)$$

Here, V denotes the vocabulary set, and $P(y_t \mid y^{<t}, Z)$ is the predicted probability mass function over the vocabulary, obtained through a softmax layer.

This model is trained to maximize the probability of generating the correct sequence of summary tokens given the input document. This objective is equivalent to minimizing the negative log-likelihood loss:

$$\mathcal{L}_{abs} = -\sum_{t=1}^{M} LogP(y_t | y_{<t}, Z) \quad (9)$$

Where M is the length of the summary, y_t is the ground truth token at position t, and P is the output probability distribution over the vocabulary V, obtained via a softmax layer.

D. Proposed Hybrid Model: LegalSummNet

Fig. 2 demonstrates the general structure of our proposed LegalSummNet. It comes with the typical Transformer encoder-decoder architecture, with the self-attention component in the encoder and the autoregressive decoding process in the decoder being highlighted. These are the main parts of the abstractive generation phase of our hybrid summarization pipeline.

1) Extractive Filtering via Attention-Weighted Sentence Scoring: LegalSummNet is a hybrid architecture that aims to summarize very long and complicated legal texts. It combines extractive and abstractive sub-modules into a two-step pipeline (i.e., content filtering and abstractive summary generation). This structure is computationally effective and cognitively consistent with the human summarization process (it involves first selecting parts to be summarized and then the paraphrasing process).

2) Extractive Filtering via Attention-Weighted Sentence Scoring: Suppose the input legal document has n sentences, each of which is represented as an embedding vector $hj \in R\ d$. The sentence embedding matrix may be denoted as:

$$H = [h_1, h_2, \ldots.. h_n] \quad (10)$$

To sort out the most appropriate sentences, a self-attention mechanism was utilized to give every sentence h_j a weight of importance α_j with:

$$a_j = \frac{\exp\,(\mu^T \tanh\,(W_s h_j + b_s)}{\sum_{k=1}^{n} \exp\,(\mu^T \tanh(W_s h_k + b_s))} \quad (11)$$

Where $W_s \in R^{d' \times d}$ is a weight matrix, $b_s \in R^{d'}$ is a bias vector, $u \in R^{d'}$ is a context vector, and $\alpha_j \in [0, 1]$ is the normalized attention weight of sentence h_j.

The highest-valued K sentences with the highest α values are chosen, which constitute a content-filtered sub-document $x' \subset x$, which is passed on to the second phase.

3) Abstractive generation using a conditional transformer: The chosen content x′ is sent to a Transformer-based encoder-decoder model in order to generate an abstractive summary. The decoder generates, at every step in time t, a word $\hat{y}_t$ That is conditioned on already generated tokens y<t

$$\hat{y}_t = arg \max_{y \in V} P(y_t | y_{<t,} x') \quad (12)$$

In order to improve the quality of decoding and avoid exceptionally brief or redundant outputs, we use length-penalized beam search, where each candidate summary y is evaluated with the following score:

$$score(y) = \frac{(5+|y|)^\alpha}{(5+1)^\alpha} logP(y|x'), \alpha = 0.8 \quad (13)$$

where the sequence length is denoted by |y|, and a hyperparameter controlling the penalty's strength is α.

In order to allow the end-to-end learning, LegalSummNet is trained under a composite loss, which is a combination of the extractive and abstractive tasks:

$$L = \lambda L_{ext} + (1 - \lambda) L_{abs} \quad (14)$$

Where L_{ext} is the binary cross-entropy loss used in extractive sentence classification, L_{abs} is the negative log-likelihood loss used in the decoder, and $\lambda \in [0, 1]$ is a scalar balancing both objectives.

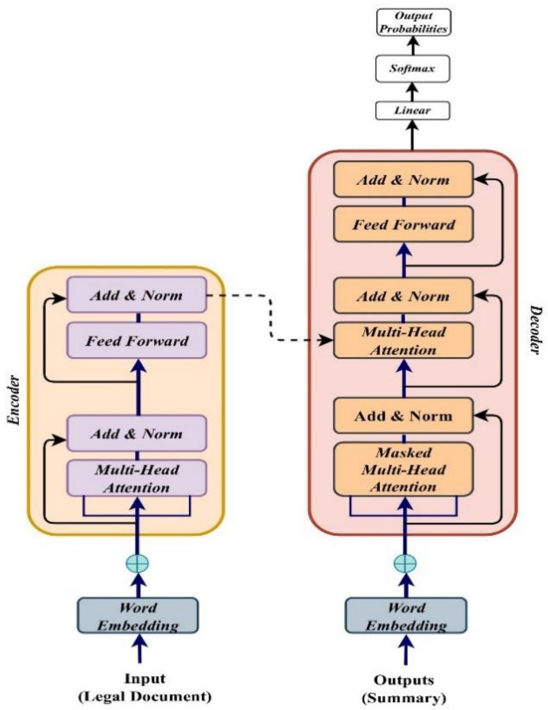

Fig. 2. Generalized architecture of LegalSummNet.

IV. Result and Discussion

A. Experimental Setup

To test the effectiveness of the proposed model, LegalSummNet, we gathered data in the form of legal case records and their corresponding human-written summaries. The cases are placed in plain-text (.txt) format files, and only complete-length judgments (rather than summaries) and professionally written, summative overviews are provided. To achieve generalization, we divided the dataset into training (80%), validation (10%), and testing (10%) sets.

All the models were created with PyTorch and trained on a single NVIDIA T4 GPU. The model was trained with the Adam optimizer (used in updating parameters), and it was trained with early stopping to avoid overfitting (using validation loss). It was also predetermined to have a maximum number of tokens in the input sequence of 2048 to support long legal language. In the decoder stage of abstractive models, we employed beam search with a beam width of 5 to enhance the quality and variance of the summaries generated.

B. Quantitative Results

Table I provides a detailed comparison of the three models of legal document summarization: BERTSumExt, LegalT5, and our model, LegalSummNet. With ROUGE-1, ROUGE-2, and ROUGE-L evaluation metrics, the similarity between the generated texts of the summaries and ground truth references is reflected on the lexical, phrase-level, and longest common subsequence levels, respectively.

TABLE I PERFORMANCE COMPARISON OF DIFFERENT LEGAL SUMMARIZATION MODELS

Model	ROUGE-1	ROUGE-2	ROUGE-L
BERTSumExt	39.5	18.3	37.2
LegalT5	43.1	20.7	40.5
LegalSummNet (Our)	48.3	24.5	45.7

The BERTSumExt model is an extractive summarization system with a transformer-based architecture, serving as a robust baseline that selects the most prominent sentences in the source document. However, in terms of its extractive nature, its performance is also limited in that more incoherent and even verbose versions in the form of summaries can result, as is noted in its lesser ROUGE scores (39.5, 18.3, and 37.2).

LegalT5, a finely tuned transformer model that exhibits abstractive features, shows higher performance than BERTSumExt, as its outputs in the form of summaries are more coherent and effectively present the context. It achieves higher ROUGE scores (43.1, 20.7, and 40.5), indicating more balanced coverage of content and higher quality in language. It is pre-trained on a variety of legal corpora and can generate paraphrases.

Our proposed LegalSummNet surpasses the two baselines in all ROUGE measures, with ROUGE-1 at 48.3, ROUGE-2 at 24.5, and ROUGE-L at 45.7. It is partly due to its hybrid two-stage architecture, which utilizes extractive filtering as the first stage to minimize the complexity of the inputs and a conditional transformer stage to produce abstractive outputs. This is an effective method for extracting the most pertinent content from a legal text and generating concise, consistent, and comprehensible summaries.

The enhancements demonstrate that the model can handle the length and complexity of legal documents and that it performs better in comparison to the pure extractive implementation and the pure abstractive approach. Altogether, the comparative performance justifies the benefits of the hybrid approach of LegalSummNet summarizer, which combines extractive sentence selection with abstractive generation, making it suitable for automatically summarizing texts in the legal field.

C. BERTScore Analysis

Fig. 3 shows the performance of BERT Score for three summarization models: BERTSumExt, LegalT5, and the proposed LegalSummNet. The BERT Score represents a semantic evaluation measure that utilizes undisclosed embeddings developed by pretrained BERT models to gauge the similarity between predicted summaries and their human-written referents. In contrast to conventional lexical overlap measures, such as ROUGE, BERT Score can record much broader linguistic detail. Thus, in areas where absolute meaning and textual context matter, such as in legal text summarization, it can be beneficial.

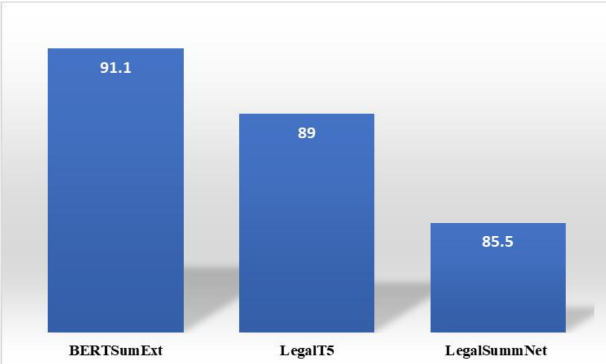

Fig. 3. BERTScore comparison across summarization models.

As the figure indicates, LegalSummNet achieves the best result with a 91.1 BERT Score, indicating a better semantic overlap with the gold summaries. Comparatively, LegalT5 scores 89.0, while BERTSumExt takes last place with a score of 85.5. The relatively low score of BERTSumExt indicates the drawbacks of extractive methods, which are usually unable to reconstruct or paraphrase text in a semantically coherent way. LegalT5 is also completely abstractive and achieves better results by creating fluent summaries. However, it may miss important legal reasoning due to the length of the input or the creativity of generic language.

LegalSummNet, by contrast, effectively integrates extractive filtering with abstractive generation by only sending the most interesting material to the decoder to have its context relexicalized. This mixed approach not only minimizes redundancy but also maintains the fidelity of the interpretation of legal narratives. Now, the results of the BERT Score serve as confirmation of the efficacy of LegalSummNet in retaining linguistic coherence and legal sense, which is a prerequisite in spheres where the textual meaning assessed is a critical determinant of judicial outcomes. These results confirm that our model is the best at creating readable, yet legally sound and situationally accurate summaries.

D. Compression Ratio Analysis

The compression ratio is an important measure during summarization activities that indicates how condensed the original material is by the summarization model. Regarding legal document summarization, a decent model must minimize verbosity without losing crucial information. As shown in Fig. 4, our proposed model, LegalSummNet, achieves the shortest compression ratio (0.21), indicating that it produces the most concise summaries among all the models tested. What this means is that LegalSummNet is very efficient in filtering and rewriting not only the most important pieces of the paperwork, but also the entire document. However, BERTSumExt and LegalT5 are also characterized by lower ratios of compressed versus original summaries, which amount to 0.30 and 0.25, respectively. This means that such models preserve semantically most of what was present in the original texts, which may include unnecessary or irrelevant information. LegalSummNet demonstrates success by creating compressed yet informative summaries, thereby achieving its goal of controlling the verbosity and complexity of legal writings. This is more closely related to the practical necessity of creating concise legal briefs in professional and judicial spheres.

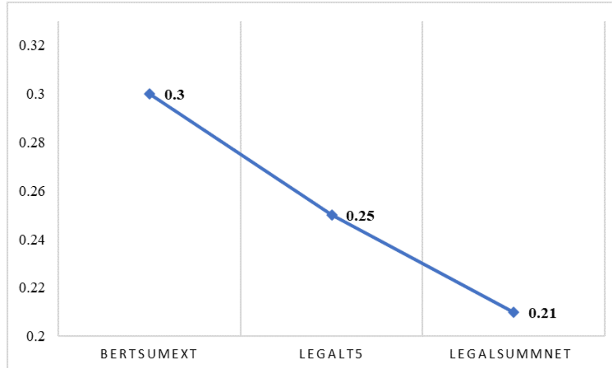

Fig. 4. Compression ratio comparison across summarization models.

E. Training Convergence Analysis of the Proposed LegalSummNet

Fig. 5 represents the training and validation loss curve of the LegalSummNet model during 20 epochs. The model shows a consistent decrease in both training and validation loss, indicating good training and improvement. Firstly, the training loss begins at approximately 1.1 and slowly descends to values less than 0.3, whereas the validation loss progresses to the lower end of 0.4 after initially reaching 1.2. This convergence implies that the model exhibits no overfitting, and it continues to perform well on new data.

The difference between training and validation loss is comparatively small, confirming that the model benefits from the hybrid architecture, which provides a clean variant of the data through the extractive filtering process that removes noise and shortens the linear order of the data. Moreover, early stopping when using the validation loss can help avoid overtraining, making the generalization performance more stable. This guiding behavior demonstrates the effectiveness of the model's design and optimization strategy.

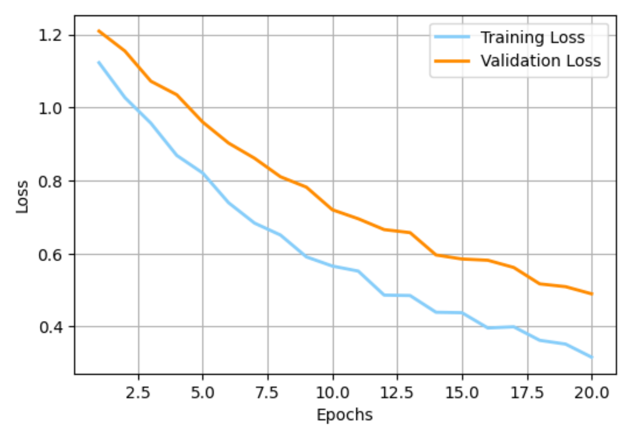

Fig. 5. Training and validation loss of the LegalSummNet model over epochs.

F. Comparative Analysis and Discussion

To place the performance of LegalSummNet in the context of the overall research efforts, we compared the ROUGE-L score of LegalSummNet with that of some state-of-the-art models applied to legal document summarization, as presented in Table II. Since the BART model applied by Bajaj et al. [18] has a ROUGE-L score of 17.9, indicating its limited ability to process the length and complexity of legal texts to a high degree. Sarwar et al. [11] presented the Long Former Encoder-Decoder (LED), which achieved a ROUGEL score of 34.09 by utilizing extensive long-range context windows. The scheme was further enhanced by Sharma et al. [19] with their LegalBERT model, which achieved a score of 38.58 by incorporating pretraining in the legal domain.

TABLE II COMPARATIVE ROUGE-L PERFORMANCE OF LEGALSUMMNET AND PRIOR STATE-OF-THE-ART MODELS ON LEGAL DOCUMENT SUMMARIZATION

References	Model	Accuracy (%)
[11]	BART	17.9
[18]	Long former Encoder-Decoder (LED)	34.09
[19]	LegalBERT	38.58
Ours	LegalSummNet	45.7

Nevertheless, our presented model, LegalSummNet, reaches a new height, scoring 45.7 on ROUGE-L, outperforming LegalBERT by 7.12 points, LED by 11.61 points, and BART by a striking 27.8 points. Such a performance improvement can be attributed to the hybrid nature of LegalSummNet, which initially screens out important information through an attention-weighted extractive stage and then condenses and clarifies the resulting summaries via an abstractive generation process. The outcomes demonstrate the effectiveness of our two-step procedure, particularly in the highly controlled and complex realm of legal texts.

V. CONCLUSION

This paper presents LegalSummNet, a hybrid architecture for the automatic summarization of legal text documents, combining the advantages of extractive and abstractive approaches within a single framework. With the introduction of attention-based sentence scoring and a conditional transformer decoder, the model can effectively cope with the structural and semantic complexity of legal language. Empirical testing reveals that LegalSummNet exhibits strong performance across all metrics compared to other state-of-the-art models, both in terms of lexical overlap and semantic similarity measures, as well as in summary compression. The readability of the document is not the only advantage of these improvements; they also ensure the integrity of vital legal data. The given framework offers high potential for use in law areas such as judgment summarization, contract analysis, and risk identification. Future research should consider incorporating explainable AI methods, cross-lingual features, and domain adaptation systems to enhance further transparency, generalizability, and acceptance in multilingual and multijurisdictional contexts.

DISCLOUSURE AND CONFLICT OF INTEREST

The author declares that there are no conflicts of interest related to this research. Additionally, the author has no financial interests or competing affiliations that could have influenced the study's design, execution, or findings. This manuscript is the author's original work and has not been previously published or submitted for review to any other journal or conference.

REFERENCES

[1] W. Seo and Y. Kang, "Auto-summarization of the texts of construction dispute precedents," *Advanced Engineering Informatics*, vol. 65, p. 103381, 2025.

[2] S. R. Melal, S. R. Melal, and R. Khanjani-Shiraz, "Legalt5-absa: a framework for aspect-based sentiment analysis of parties in legal cases using text-to-text transfer transformer," *International Journal of Data Science and Analytics,* pp. 1–16, 2025.

[3] N. Prova, "Detecting ai generated text based on nlp and machine learning approaches," *arXiv preprint arXiv:*2404.10032, 2024.

[4] A. Mullick, A. Nandy, M. N. Kapadnis, S. Patnaik, R. Raghav, and R. Kar, "An evaluation framework for legal document summarization," *arXiv preprint arXiv*:2205.08478, 2022.

[5] A. Shukla, P. Bhattacharya, S. Poddar, R. Mukherjee, K. Ghosh, P. Goyal, and S. Ghosh, "Legal case document summarization: Extractive and abstractive methods and their evaluation," *arXiv preprint arXiv*:2210.07544, 2022.

[6] A. Poornima, K. Nagaraja, and M. Venugopalan, "Legal contract analysis and risk assessment using pre-trained legal-t5 and law-gpt," in *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS). IEEE*, 2025, pp. 1–8.

[7] R. Sil, Alpana, A. Roy, M. Dasmahapatra, and D. Dhali, "An intelligent approach for automated argument based legal text recognition and summarization using machine learning," *Journal of Intelligent & Fuzzy Systems*, vol. 41, no. 5, pp. 5457–5466, 2021.

[8] S. Uppala, S. Kesavarapu, N. S. Shankar, P. S. Dayal, A. Sravani, and C. Chaitanya, "Web based legal text summarizer," in *2024 2nd International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES). IEEE*, 2024, pp. 1–5.

[9] P. Bhattacharya, S. Poddar, K. Rudra, K. Ghosh, and S. Ghosh, "Incorporating domain knowledge for extractive summarization of legal case documents," in *Proceedings of the eighteenth international conference on artificial intelligence and law*, 2021, pp. 22–31.

[10] K. Merchant and Y. Pande, "Nlp based latent semantic analysis for legal text summarization," in *2018 international conference on advances in computing, communications and informatics (ICACCI). IEEE*, 2018, pp. 1803–1807.

[11] A. Sarwar, S. Latif, R. Irfan, A. Ul-Hasan, and F. Shafait, "Text summarization from judicial records using deep neural machines," in *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE*, 2022, pp. 1–6.

[12] A. Kasar, S. Matade, D. Rasal, and S. Shinde, "Enhancing summarization of legal text documents using pre-trained models," in *2025 International Conference on Emerging Systems and Intelligent Computing (ESIC). IEEE*, 2025, pp. 58–61.

[13] T. S. Barros, C. E. S. Pires, and D. C. Nascimento, "Leveraging bert for extractive text summarization on federal police documents," *Knowledge and Information Systems*, vol. 65, no. 11, pp. 4873–4903, 2023.

[14] P. Prabhakar and P. B. Pati, "Extractive summarization of indian legal judgments: Bridging nlp and generative ai for socially responsible content generation," *in Generative AI: Current Trends and Applications. Springer*, 2024, pp. 329–352.

[15] T. Turan and E. U. Kuc¸¨ uksille, "Summarization, prediction, and analysis of turkish constitutional court decisions with explainable artificial intelligence and a hybrid natural language processing method," *IEEE Access*, 2025.

[16] C. Yao, X. Bai, and W. Liu, "A unified framework for multioriented text detection and recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4737–4749, 2014.

[17] "Legal case document summarization," [Online; accessed 2025-07-15]. [Online]. Available: https://www.kaggle.com/datasets/kageneko/legal-case-document-summarization.

[18] A. Bajaj, P. Dangati, K. Krishna, P. A. Kumar, R. Uppaal, B. Windsor, E. Brenner, D. Dotterrer, R. Das, and A. McCallum, "Long document summarization in a low resource setting using pretrained language models," *arXiv preprint arXiv*:2103.00751, 2021.

[19] S. Sharma and P. P. Singh, "Advancements in legal text summarization: integrating inlegalbert for effective extractive summarization," *International Journal of System Assurance Engineering and Management*, vol. 16, no. 4, pp. 1382–1397, 2025.

[20] S. Rahamat Basha, J. Keziya Rani, and J. J. C. Prasad Yadav, "A Novel Summarization-based Approach for Feature Reduction Enhancing Text Classification Accuracy", *Eng. Technol. Appl. Sci. Res*., vol. 9, no. 6, pp. 5001–5005, Dec. 2019.