

# Large Language Models for Indian Legal Text Summarisation

(1571006931)



Hemanth Kumar M, Jayanth P and Anand Kumar M  
Artificial Intelligence, Department of Information Technology  
National Institute of Technology Karnataka  
Surathkal, India

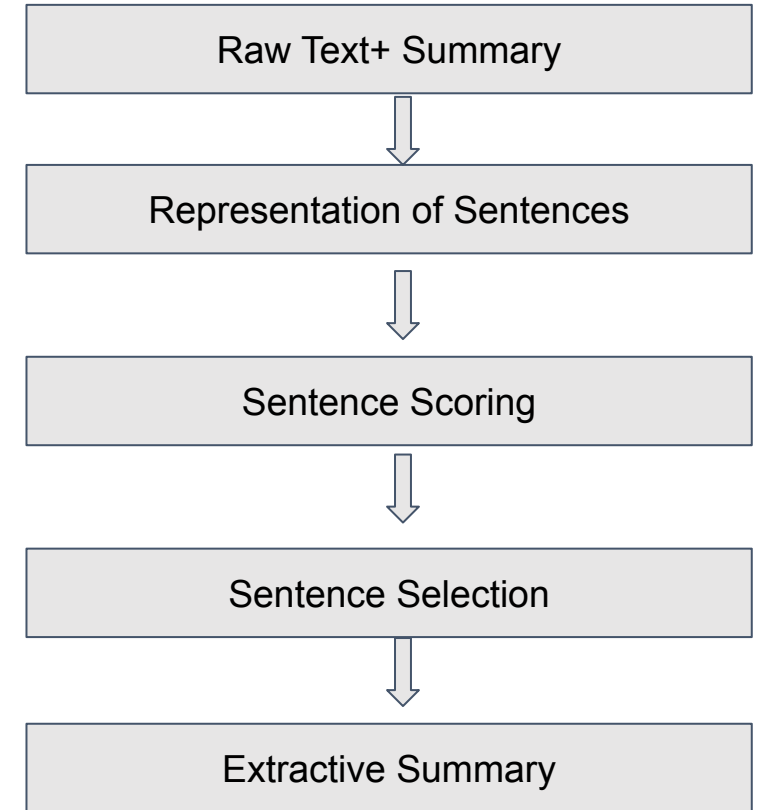
# Agenda

- Introduction
- Objectives
- Methodology
- Results and Discussion
- Conclusion
- References

# Introduction

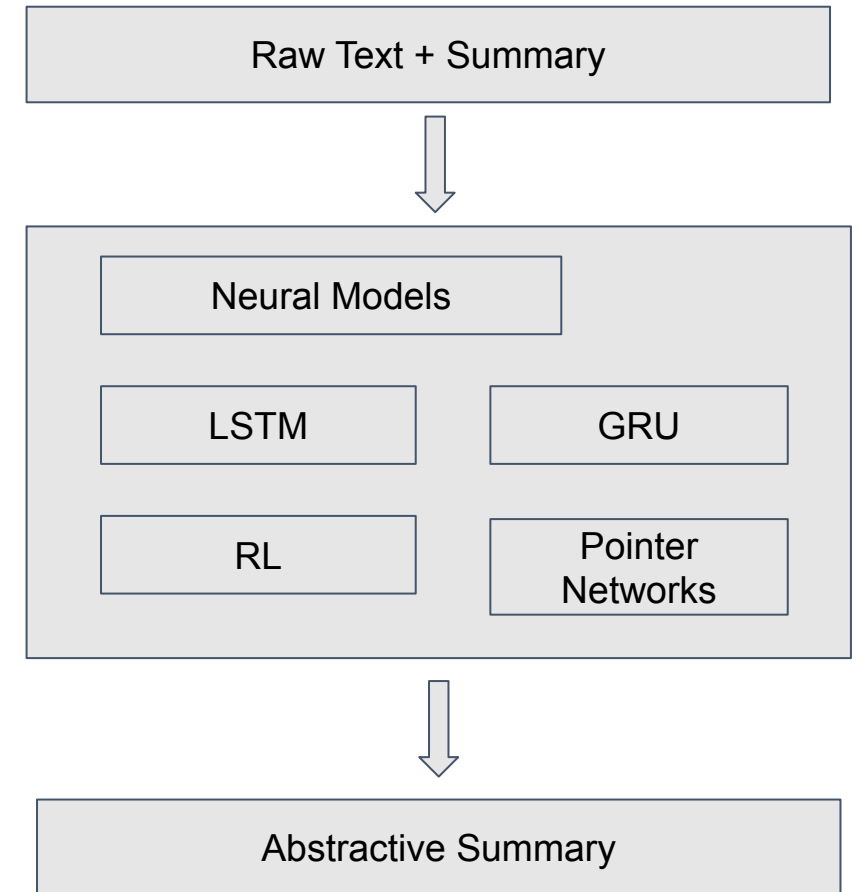
There are two main approaches to summarization:

**Extractive Summarization** identifies important sentences or phrases from the text and extracts them to create a summary. Nevertheless, it occasionally results in summaries that may seem fragmented or disconnected.



# Introduction

**Abstractive Summarization** entails creating a novel summary that represents a fluid and coherent rephrasing of the original text.



# Introduction...

- Document summarization is not a new problem. There are many document summarizations like:
  - News document summarization
  - Scientific article summarization.
- But, there is a difference in Legal document summarization. A court case or legal document is long and there are no segments and no para headings. This makes the legal document summarization bit more complicated.

## Challenges



# Introduction...

- Legal documents, known for their length and unique abbreviations, often require labor-intensive manual summarization.
- The emergence of Artificial Intelligence (AI) and Machine Learning (ML) has paved the way for automatic summarization, potentially saving significant time and effort [4].
- This innovation holds particular promise in the legal domain, not only benefiting professionals but also facilitating understanding for beginners and the general public.
- With an overwhelming backlog of over 4.70 crore pending cases in various Indian courts [6], the adoption of automatic summarization came as a potential solution.

# Objectives

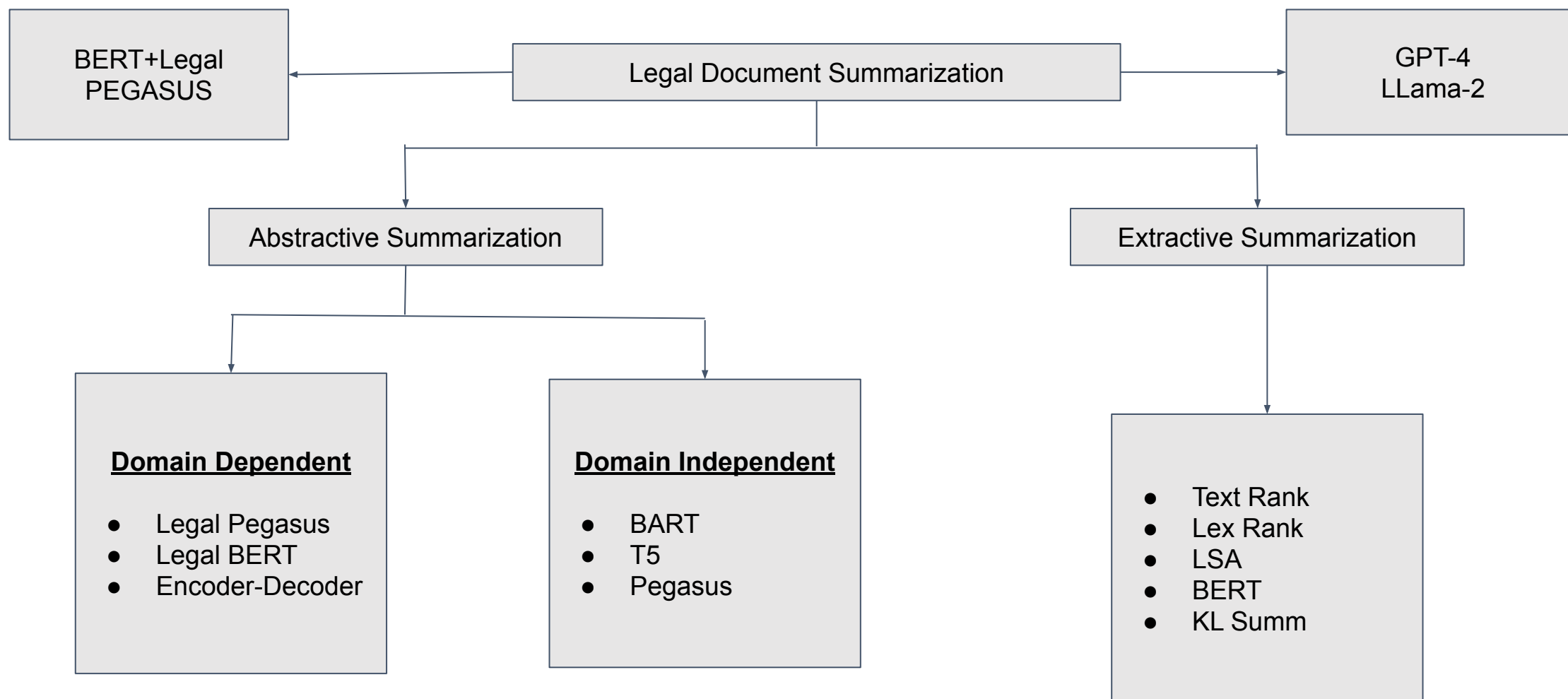
- To build a Hybrid Approach for legal document summarization. As the legal documents are large , if Hybrid approach is used, it can reduce the complexity of the document.
- While extensive literature exists on extractive summarization, a notable gap is still maintained in the domain of Abstractive Summarization, which increased research attention.
- The literature survey highlights a substantial research gap in the development of evaluation metrics tailored to the legal domain for summarization tasks . Existing metrics such as ROUGE are deemed insufficient in gauging the true quality of system generated summaries [5]. In response, leveraging BERT for paragraph similarity calculation emerges as a viable approach to address existing limitations.

# Dataset

- A scarcity of publicly available datasets specifically designed for legal case document summarization, particularly in the English language, is evident in the current landscape.
- The dataset [3] , we collected contains 7130 documents for Abstractive summarisation. Out of which, 5624 are used for training and 1406 are used as validation, the remaining 100 documents are used for testing.
- There are 50 documents for Extractive summarisation task.



# Methodology



# Results & Discussion

Model	ROUGE-1	ROUGE-2	ROUGE-L	BERT SCORE
<b>BART</b>	0.32	0.09	0.18	0.69
<b>T5</b>	0.12	0.02	0.093	0.57
<b>PEGASUS</b>	0.30	0.06	0.17	0.75
<b>RoBERTa</b>	0.34	0.04	0.19	0.64
<b>Legal PEGASUS</b>	0.43	0.12	0.27	0.77
<b>Legal BERT</b>	0.19	0.05	0.27	0.62
<b>Encoder - Decoder</b>	0.21	0.08	0.17	0.68

Performance of the models for Abstractive summarization

# Results & Discussion

Model	ROUGE-1	ROUGE-2	ROUGE-L	BERT SCORE
LSA	0.30	0.07	0.14	0.71
Lex Rank	0.32	0.11	0.16	0.82
Text Rank	0.44	0.12	0.18	0.85
KL Summ	0.20	0.06	0.19	0.63
BERT	0.28	0.07	0.14	0.59

Performance of the models for Extractive summarization

Model	ROUGE-1	ROUGE-2	ROUGE-L	BERT SCORE
LongFormer	0.24	0.11	0.19	0.81
Hybrid Model	0.25	0.07	0.14	0.73

Performance of Hybrid models for the task Legal Text summarization

# Results & Discussion

Metric	Zero-shot	One-shot
ROUGE-1	0.301	0.228
ROUGE-2	0.098	0.066
ROUGE-L	0.250	0.179
BERT SCORE	0.779	0.795

Performance of LLama for the task Legal Text summarisation (Abstractive)

Metric	Zero-Shot	One-Shot
ROUGE-1	0.368	0.335
ROUGE-2	0.115	0.048
ROUGE-L	0.216	0.176
BERT SCORE	0.65	0.68

Performance of LLama for the task Legal Text summarization(Extractive)

# Results & Discussion

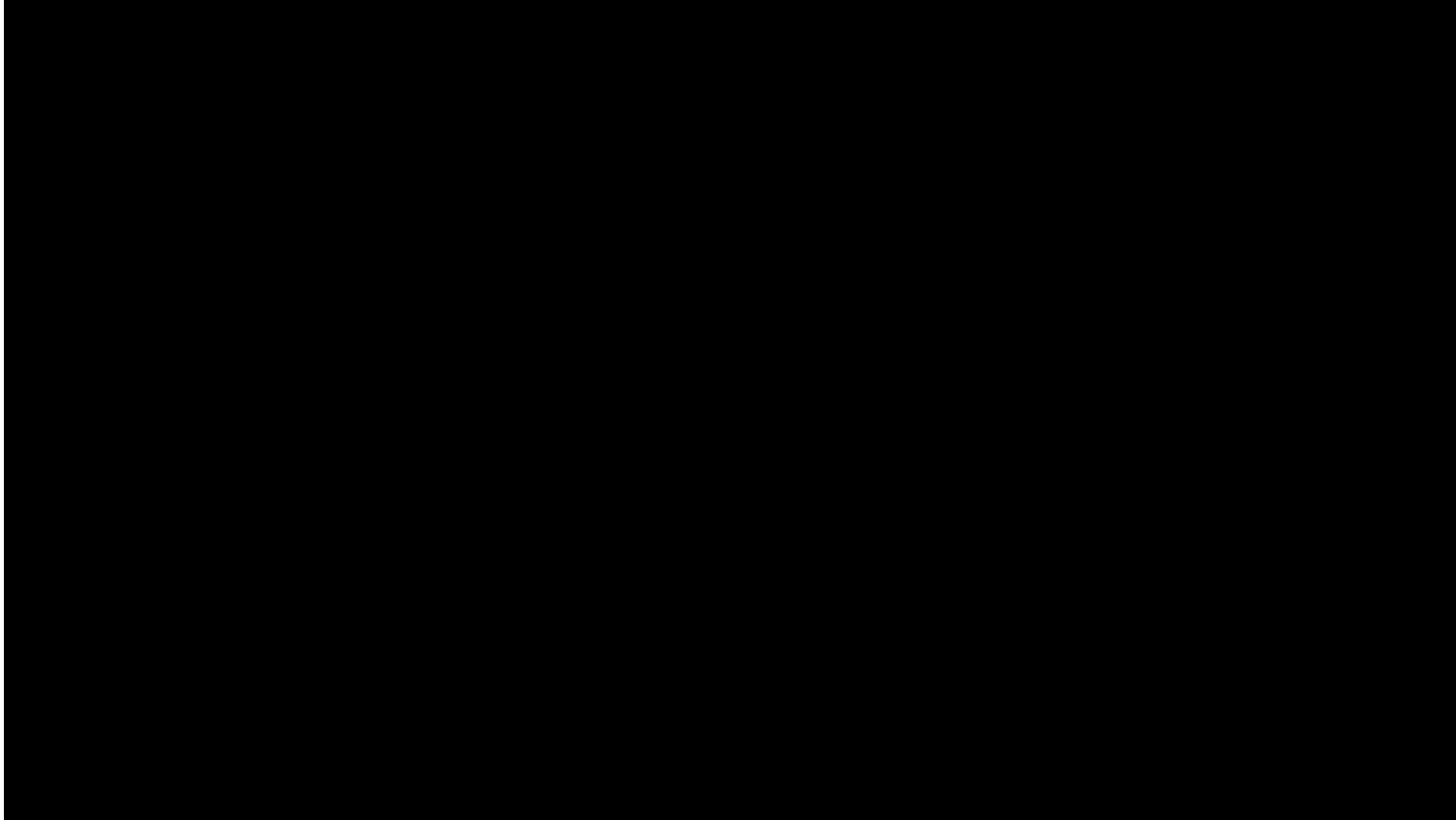
Metric	Zero-shot	One-shot
ROUGE-1	0.18	0.31
ROUGE-2	0.05	0.12
ROUGE-L	0.16	0.29
BERT SCORE	0.78	0.82

Performance of GPT-4 for the task Legal Text summarization (Abstractive)

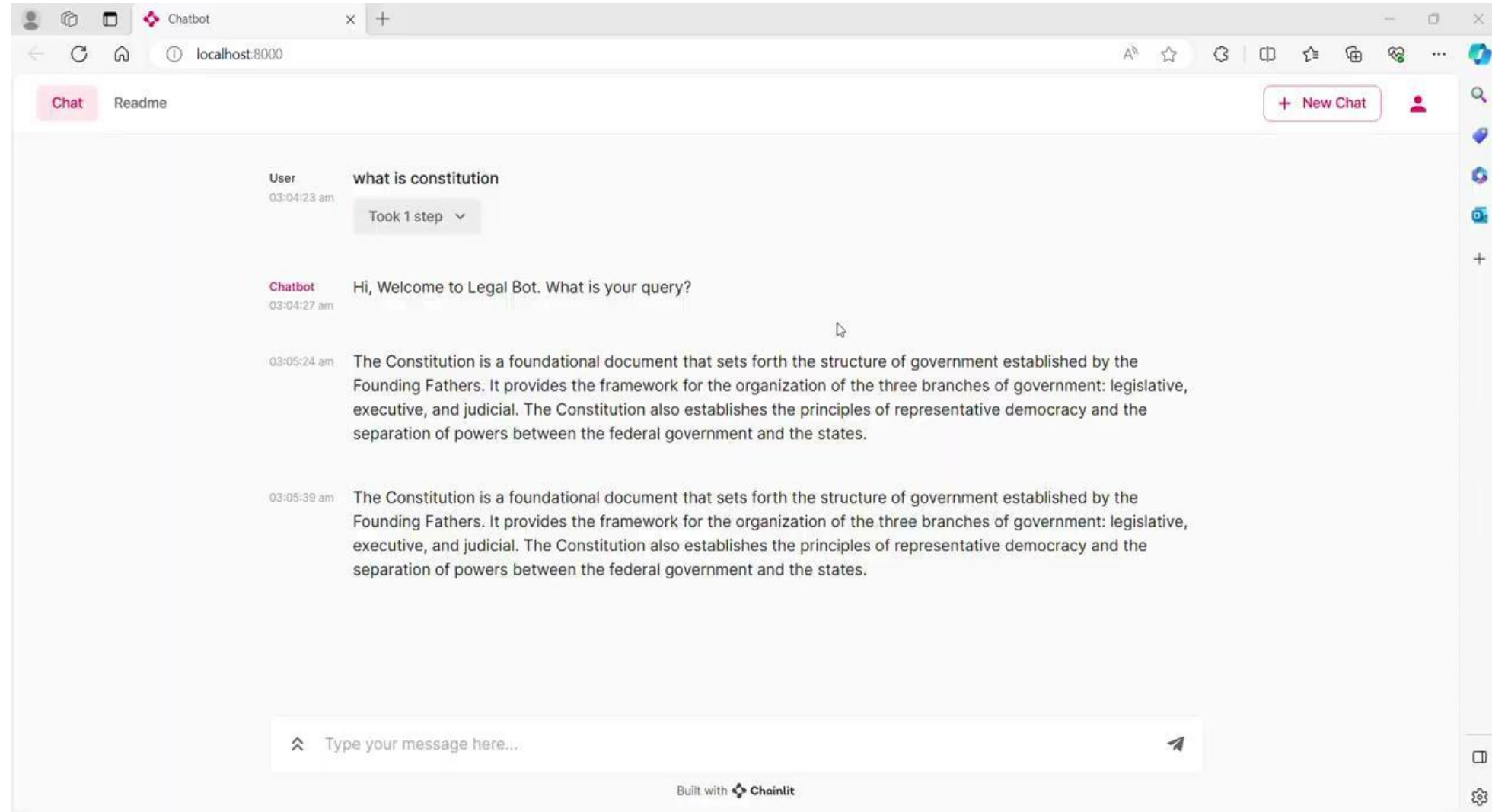
Metric	Zero-Shot	One-Shot
ROUGE-1	0.31	0.51
ROUGE-2	0.047	0.105
ROUGE-L	0.167	0.217
BERT SCORE	0.69	0.79

Performance of GPT-4 for the task Legal Text summarization(Extractive)

# Results & Discussion



# Results & Discussion



# Conclusion

- In summary, our work in legal text summarization introduced a novel approach in Indian legal context for both extractive and abstractive, using BART, T5, Pegasus, Roberta models and few legal domain specific models such as Legal-Pegasus, Legal-BERT.
- We Considered our work with previous works, few considered different datasets.
- Our dataset is the part of original dataset because of resource constraints. We can't compare our results with the original paper results.
- As for our knowledge, our's is the first work that implemented LLM like llama-2 and GPT-4 for Legal document summarisation.
- The proposed Legal Pegasus model obtained best results via rouge as well as bert score in abstractive summarization because of the advantage of pretraining in Legal domain.
- Hybrid Longformer methodology significantly improves ROUGE scores, emphasizing the effectiveness of domain specific adaptations.
- As a future scope, we aspire to look into the domain of summarizing legal text by employing instruction fine-tuning along with large language models, such as Llama-2 and GPT-4.



# References

- [1] Surden, Harry. "Artificial intelligence and law: An overview." Georgia State University Law Review 35 (2019): 19-22.
- [2] PTI, "Over 4.70 crore cases pending in various courts: Govt," 3 2022.
- [3] <https://zenodo.org/records/7152317.Yz6mJ9JByC0>
- [4] 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.
- [5] Zhang, Tianyi, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. "Bertscore: Evaluating text generation with bert." arXiv preprint arXiv:1904.09675 (2019).
- [6] Anand, Deepa, and Rupali Wagh. "Effective deep learning approaches for summarization of legal texts." Journal of King Saud University-Computer and Information Sciences 34.5 (2022): 2141-2150.

Thank You

Q&A