

PROJECT PROPOSAL

LegalEase: Summarization of legal documents and query-answering chatbot

Kunal (2021330)

Ishan (2021465)

Aman Ranjan (2021376)

Megha (2021337)

6th February, 2024

Problem Statement

It takes a lot of time and confusion to sort through the complex legal documents; The process is irritating. In particular, when it comes to contracts, loans, and real estate transactions, current solutions fall short of providing user-friendly experiences. This lack of simplicity and clarity is a significant obstacle to making confident, well-informed decisions in essential transactions.

- Time-consuming
- Confusing
- Contracts
- Unapproachable
- User-friendly
- Lack of simplicity
- Obstacle
- Intimidating
- Loans
- Real estate
- Informed decisions
- Lack of clarity

Importance of the Problem

Most of us have faced issues related to this problem. Let's see what Kunal faced

“A few years ago, my family bought an apartment. We were told that everyone would share the terrace. This was important to us. Later, the builders decided to add another floor and give the terrace to the top floor. When we complained, they showed us the agreement. It was 50 pages long and we needed to read it all carefully. It said that the builders could make these changes. We were disappointed but couldn't do anything.

This experience made us realize how hard it is to understand legal documents.”

A wide range of people are impacted by the challenge of understanding legal documents, including small businesses without dedicated legal departments and individuals engaged in personal transactions. Failing to comprehend the terms and conditions fully can have negative consequences, such as monetary losses and legal disputes. As a result, making it easier to understand these documents is vital for financial stability, legal safety, and convenience.

Related Work

Several services and tools, such as chatbots that provide legal advice and document summarization tools, aim to demystify legal documents. Nevertheless, these solutions frequently need to offer interactive, document-specific guidance and customized, thorough summaries. Most currently available tools are generic and need to adjust to the particular complexities of individual documents. Moreover, they change the real meaning of the legal thing.

Some students have worked on documents related to the judgment of Indian courts and made a summarization tool. However, we would like to train better in another legal domain and integrate a chatbot.

Proposed Solution and Novelty

LegalEase is an innovative tool designed to simplify complex legal documents. It transforms contracts, loans, and real estate agreements into concise, easy-to-understand summaries. Users receive a clear breakdown of the key points and terms by uploading PDFs or scanning images of legal documents. We will also create an interactive chatbot tailored to each document, allowing users to ask specific questions and receive instant, understandable answers. This approach saves time and gives users confidence and clarity. In the future, we also plan to integrate it using a website extension.

Novelty: Existing products like AI, GPTs, and others are primarily trained in foreign databases. But in the Indian context there are less similar things. And also, there will be an option for multi-language support. People also sign on to various T&Cs online, a website extension will also be provided to take care of scams.

Technologies

Using information retrieval, we will build LegalEase. These techniques will include:

- Document summary algorithms to extract essential terms and points.
- Machine learning models that can comprehend and respond to user inquiries about the document in chatbot encounters. We are thinking of using torch, BERT, transformers and similar things
- Optical character recognition (OCR) technology creates machine-readable text from scanned documents and photos.
- A front-end app or a website.
- Creation of website extension.

Evaluation

Our evaluation approach includes

- user studies that assess time savings and comprehension improvements.
- There are some methods listed in the research paper listed below.

- Usability testing is necessary to ensure the interface is accessible and easy to use.
- Many machine learning techniques can also be used to check the efficiency.

Contributions

- Each 4 of us will be involved in Model Training.
- Research papers were divided equally.
- Designing an app or a website and linking backend with front-end
- Website extension

Literature Review

1.Indian Legal Text Summarization: A Text Normalization-based Approach: Satyajit Ghosh, Mousumi Dutta, and Tanaya Das' research work addresses the difficulty of legal text summary in India. There are over 4 million pending cases. Their research offers a text normalization-based approach for improving the performance of domain-independent machine learning models such as BART and PEGASUS. Even without domain-specific training data, their experiments reveal that summarization efficacy improves significantly.

<https://arxiv.org/pdf/2206.06238.pdf>

2.The Right to remain plain : Summarization and simplification of legal : This stanford paper discusses about the existing work in the domain of text summarization and how to build upon these technologies using fine tuning within the legal domain.Also it further opens the possibilities of training the models on more specific domains in law. It uses BART as an example.

[Stanford Paper](#)


Document

3.Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation:

The paper written by Abhay Shukla and others addresses challenges in legal case summarization.It has two models listed which is comparing extractive and abstractive models. It introduces three legal summarization datasets, conducts extensive experiments.It will be really helpful in testing our model .

<https://aclanthology.org/2022.aacl-main.77.pdf>

4. Summarization of Legal Documents: Where are We Now and the Way Forward : The paper explores different ways to summarize legal stuff, focusing on how these methods work in the legal world,moreover tools that help lawyers and gives examples of how summarizing can help in



real legal situations. The research raises several important issues, such as how changes in legal document structuring in various domains affect how well they may be summarized

<https://www.sciencedirect.com/science/article/abs/pii/S1574013721000289>

5. Extracting the Gist of Chinese Judgments of the Supreme Court:

In this the author has discussed about the dataset summarisation in Chinese context and the techniques being used there. In order to generate legal gist statements from Supreme Court rulings, the study classifies sentences using machine learning techniques, such as gradient boosting and deep learning with LSTM. Gradually they were able to achieve a great F1 score.

<https://dl.acm.org/doi/10.1145/3322640.3326715>