



L.U.F.Y.

Law Understandable For You



Team PARADOX

Firoz Uddin Gauhar
Mohammad Ahmadullah Khan
Hammad Ali
Aabid Hussain

CONTENTS

1. Introduction

- Overview
- Problem Statement
- Solution
- Key Objectives

2. Idea Description

- Innovation and Uniqueness
- Value Proposition

3. Technology Overview

- Technology Stack
- Scalability and Performance

4. Impact Assessment

- Societal Impacts

5. Conclusion

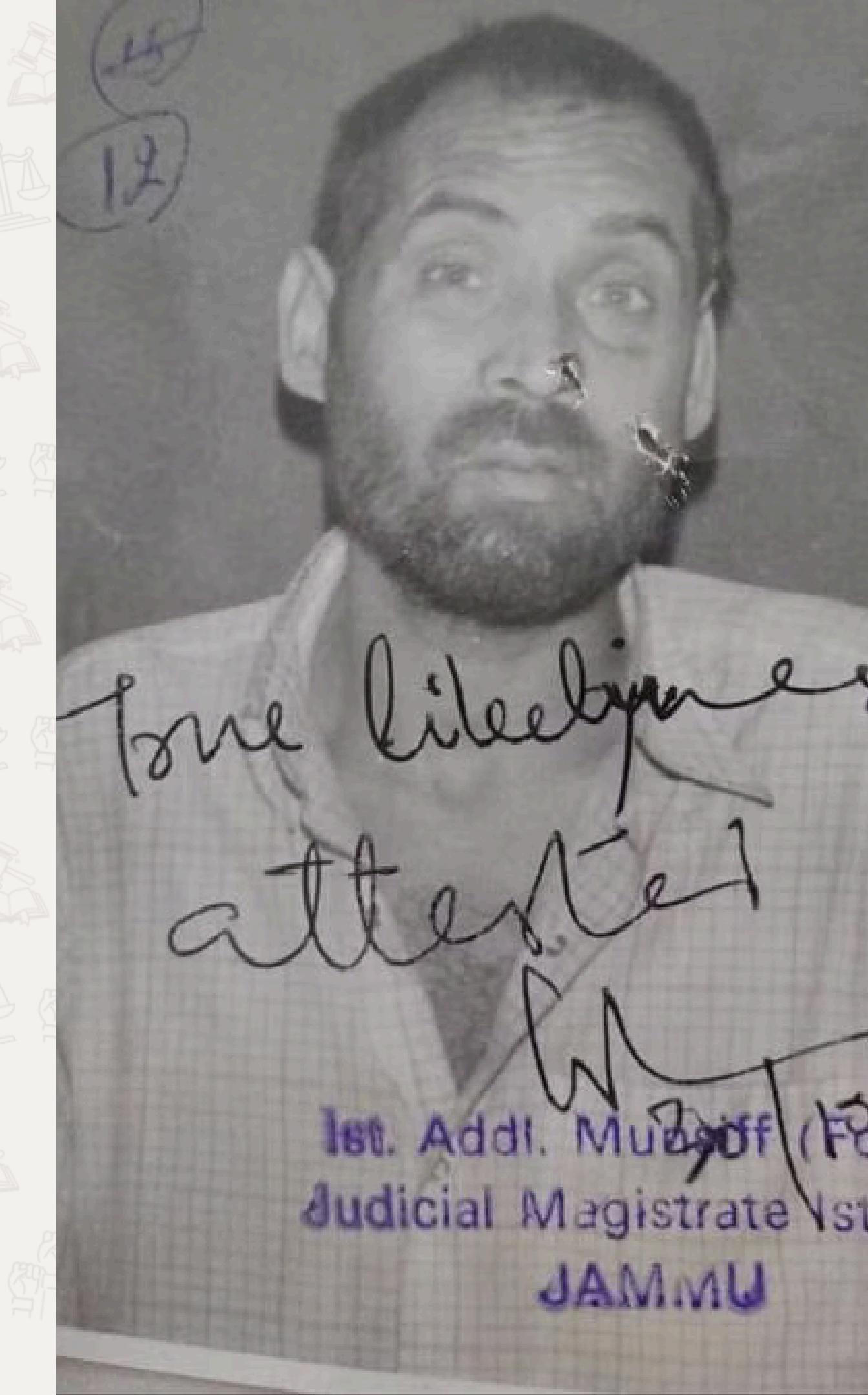


OVERVIEW

THE ORDEAL OF AN COMMON MAN IN INDIAN LEGAL SYSTEM

Imagine a world where justice isn't shrouded in legalese! In India, the legal system feels like a giant library with locked doors. Most folks stare at these documents filled with labyrinthine sentences and words straight out of a history textbook. It's like trying to decipher a secret code! This creates a system that's rigged against the common man. We can't fight for fair treatment in our rents, understand our employment contracts, or even navigate a basic police report if we can't even understand the rulebook!

Think about it: how can a street vendor understand their rights when a simple eviction notice reads like a Shakespearean sonnet? How can a young woman navigate a harassment complaint if the legalese sounds like a different language altogether? This isn't just about frustration – it's about a fundamental lack of access to justice. We deserve a legal system that speaks to us, not down to us. We deserve to understand the laws that govern our lives!



PROBLEM STATEMENT

RHETORICAL LABELING AND SUMMARIZATION OF LEGAL DOCUMENTS

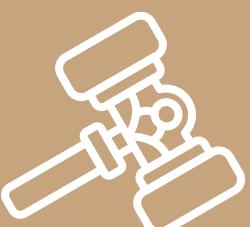
SOLUTION

In an effort to bridge the gap between the complexities of the legal system and the everyday person, we propose the development of a legal judgment summarization system. This innovative technology leverages the power of artificial intelligence to empower citizens by creating clear, concise summaries of legal documents.

This is like a magic decoder ring for the law! This tech can break down those monstrous legal documents into bite-sized, clear summaries. Imagine – eviction notices explained in the same language you'd use with your friend, contracts written as clearly as an online shopping agreement, and police reports that explain your rights like a helpful neighbor. This isn't just about convenience – it's about leveling the playing field. With clear summaries, everyone can understand their rights, navigate legal processes with confidence, not fear.

Think about a small business owner who can finally decipher that complex loan agreement without a lawyer's help. Imagine a tenant facing eviction who can now understand their rights and fight for them in court, summary in hand. This system can empower a young couple to understand their prenuptial agreement, ensuring both parties are protected. It's about removing the fear of the unknown and replacing it with the power of knowledge.

Welcome to a future where law meets simplicity and accessibility.
LUFY - Your Bridge to Legal Confidence.



SOLUTION (TECHINICAL STEPS)

Model Training:

- Train two separate models:
 - Google T5-small: A small, lightweight language model from Google AI.
 - Facebook BART-large-cnn: A large, complex language model from Facebook AI Research.
 -

Model Comparison:

- Evaluate the performance of both models on a benchmark dataset of legal judgments.
- Select the best-performing model for further development.

Dataset:

- Utilize the dataset for summarising and another dataset for rhetorical labelling.

API Development:

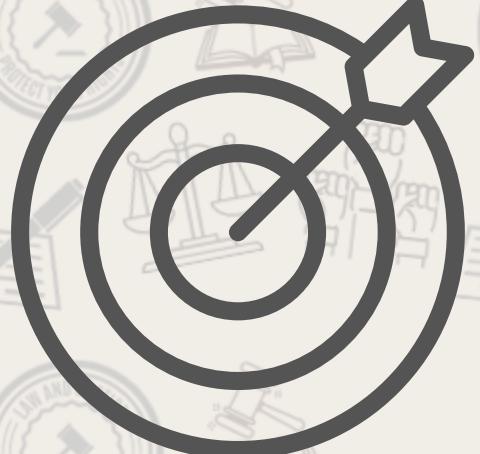
- Employ Flask, a Python web framework, to create an API for the selected model.
- The API should allow users to submit legal judgments and receive summaries in response.

Chatbot Development:

- Develop a chatbot using the API to provide a user-friendly interface for legal judgment summarization.
- Users should be able to upload documents or enter text directly into the chatbot to obtain summaries.



KEY OBJECTIVES



Enhance Public Legal Literacy

This project translates complex legal jargon into clear summaries, delivered in your own language. By breaking down these barriers, everyone can understand their rights and navigate the legal system with confidence.

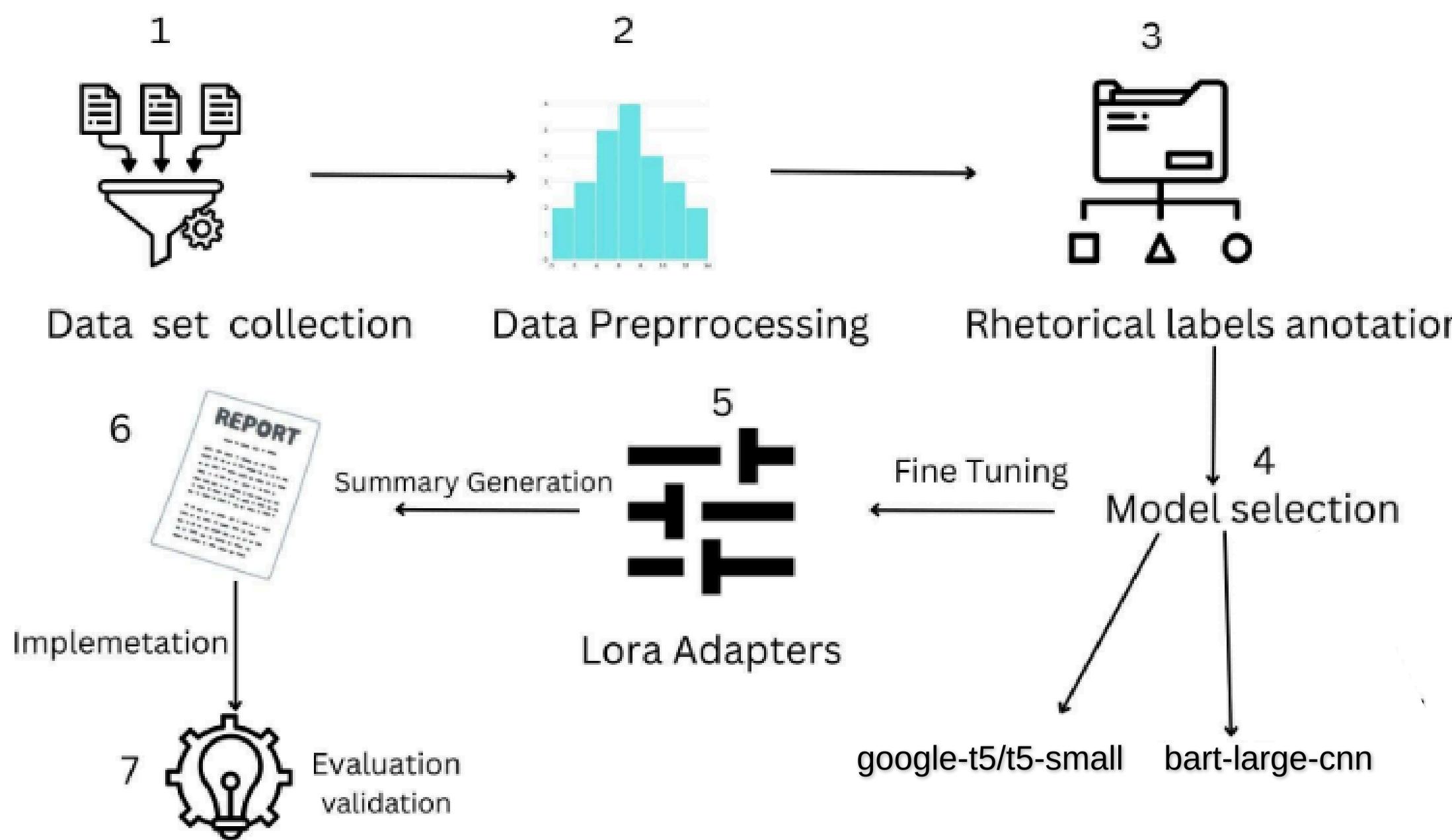
Promote Transparency and Accountability

Legal jargon often obscures court decisions. This project tackles that by offering clear summaries in local languages. By empowering citizens to understand court rulings, the project fosters transparency, leading to a more informed public that can hold the legal system accountable.

Boost Legal Access

For legal professionals, summaries expedite case review, improving efficiency for judges and lawyers who can grasp the gist of cases much faster. This translates to a more streamlined legal system for all.

SOLUTION (TECHINICAL STEPS)



SOLUTION (TECHNICAL ASPECT)

DATASET USED FOR SUMMARISING

To build a comprehensive pre-training corpus, we collected a vast corpus of legal texts from multiple sources. Specifically, we gathered case documents from the Indian Supreme Court and various High Courts across India.

The sources included:

- **Official Websites:** Documents were downloaded from official court websites, such as the Indian Supreme Court's website (<https://main.sci.gov.in/>), which ensures the data's authenticity and reliability.
- **Legal Information Institute of India:** Although the original website is no longer active, we retrieved archived case documents from this repository.
- **Indian Kanoon:** We leveraged the popular legal repository (<https://www.indiankanoon.org>) to access a wide range of case documents covering different legal topics and jurisdictions.

This collection process resulted in a dataset comprising over 50,000 legal documents, covering diverse cases and legal issues, which provided a robust foundation for training and fine-tuning our models.

SOLUTION (TECHNICAL ASPECT)

PERFORMANCE OF THE TWO TRAINED MODELS

hammadali1805/legal_bart_large_cnn

PEFT TensorBoard Safetensors Generated from Trainer License: mit

Model card Files and versions Training metrics Community

legal_bart_large_cnn

This model is a fine-tuned version of [facebook/bart-large-cnn](#) on an unknown dataset. It achieves the following results on the evaluation set:

- Loss: 2.2520
- Rouge1: 0.489
- Rouge2: 0.2433
- Rougel: 0.3262
- Rougelsum: 0.3267
- Gen Len: 137.2973

hammadali1805/model

PEFT PyTorch TensorBoard Safetensors Generated from Trainer License: apache-2.0

Model card Files and versions Training metrics Community

? model

This model is a fine-tuned version of [google-t5/t5-small](#) on an unknown dataset. It achieves the following results on the evaluation set:

- Loss: 2.3571
- Rouge1: 0.1552
- Rouge2: 0.0673
- Rougel: 0.1307
- Rougelsum: 0.1308
- Gen Len: 18.9893

using facebook/bart-large-cnn

using google-t5/t5-small

SOLUTION (TECHNICAL ASPECT)

EXAMPLE

In [3]:

```
# Example usage
text = """
petition (Crl.) No. 8193A of 1981. (Under Article 32 of the Constitution of India) R. K Garg, V. J. Francis, Sunil Kumar Jain
Dismissing the petition, the Court ^ HELD: 1. Whatever be the guidelines observed for the exercise of the power conferred by
"""
summary = summarize(text)
print("Summary:", summary)
```

Summary: The President has the power in an appropriate case to commute any sentence imposed by a court into a lesser sentence. The question as to whether the case is appropriate for the exercise of the power conferred by article 72 depends upon the facts and circumstances of each particular case. The necessity or the justification for exercising that power has therefore to be judged from case to case. We do not see what useful purpose will be achieved by the petitioner by ensuring the imposition of any severe, judicially evolved constraints on the sentence of death imposed upon him.

INNOVATION AND UNIQUENESS

SYNERGY OF LLMs AND RHETORICAL LABELS

Unlike traditional summarization methods, our system leverages the power of advanced Large Language Models (LLMs) in tandem with rhetorical labels. These labels act as a roadmap, allowing the system to differentiate between factual information, legal arguments, and conclusions within a judgment. This nuanced understanding leads to summaries that capture the logical flow and critical reasoning present in legal documents, a significant advancement over current methods.

FOCUS ON LEGAL SPECIFICITY

Existing summarization systems are often generic and struggle with the unique complexities of legal language. Our system addresses this by incorporating domain-specific training data and fine-tuning LLMs on vast amounts of legal text. This ensures the system can accurately grasp legal terminology and concepts, resulting in summaries that are not only concise but also legally sound.

MULTILINGUAL ACCESSIBILITY

Recognizing the diverse linguistic landscape in India, our system prioritizes accessibility by offering summaries in various local languages. This breaks down language barriers and empowers individuals across the country to understand legal judgments that impact their lives, and fostering a truly inclusive legal landscape for all.

VALUE PROPOSITION

This legal judgment summarization system offers a compelling value proposition for various stakeholders:

1. CITIZENS:



Gain a clear understanding of their legal rights and navigate legal processes with newfound confidence. Summaries in local languages eliminate language barriers and empower individuals to participate meaningfully in the legal system.

2. Legal Professionals:



Lawyers and judges can leverage summaries to expedite case review and enhance overall efficiency. By grasping the gist of cases more quickly, they can dedicate more time to in-depth analysis and strategic decision-making.

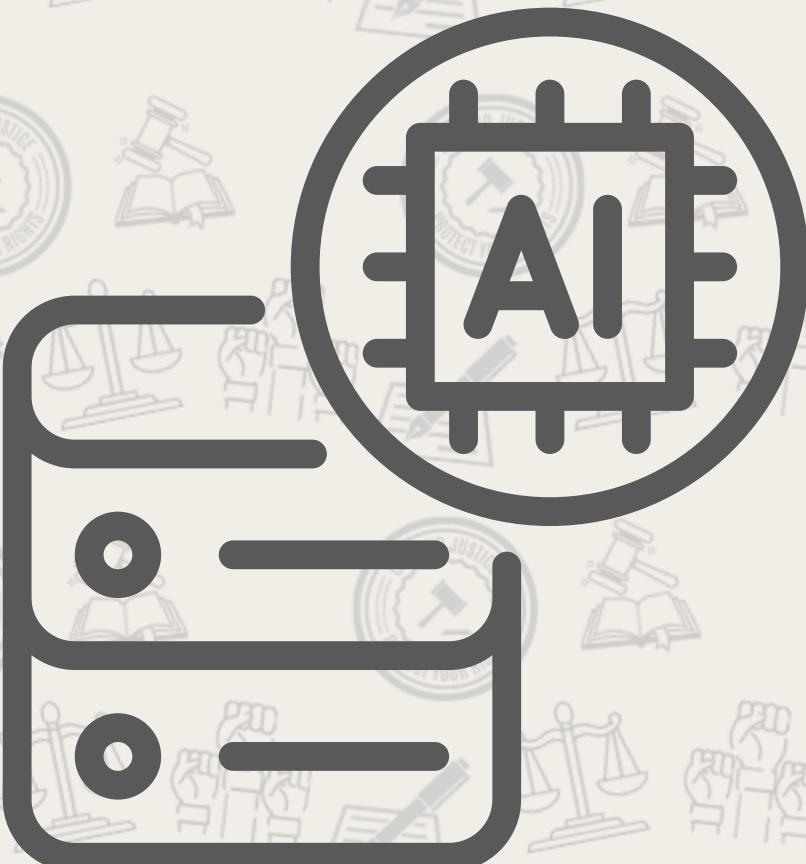
3. Justice System:



Increased public understanding of legal proceedings fosters greater transparency and accountability. This promotes trust in the legal system and empowers individuals to hold institutions accountable for upholding the law.

TECH STACK

- Python 3.8: The primary programming language used for implementation.
- PyTorch 1.10: Used for model development and fine-tuning, given its flexibility and extensive support for deep learning.
- Hugging Face Transformers: Utilized for accessing and customizing pre-trained models such as Llama 2, Legal BERT, and Legal-RoBERTa.
- Scikit-learn: Employed for preprocessing tasks and performance evaluation.
- Jupyter Notebooks: Used for interactive development and result visualization



SCALABILITY AND PERFORMANCE

- **Leveraging Advanced LLMs:**

While the initial model might utilize effective LLMs like Legal-BERT or Legal-RoBERTa, we can explore integrating the latest advancements like GPT-3 or LaMDA (Language Model for Dialogue Applications) in future iterations. These models possess even greater capabilities in understanding complex language and generating nuanced summaries.

- **Domain-Specific Fine-Tuning:**

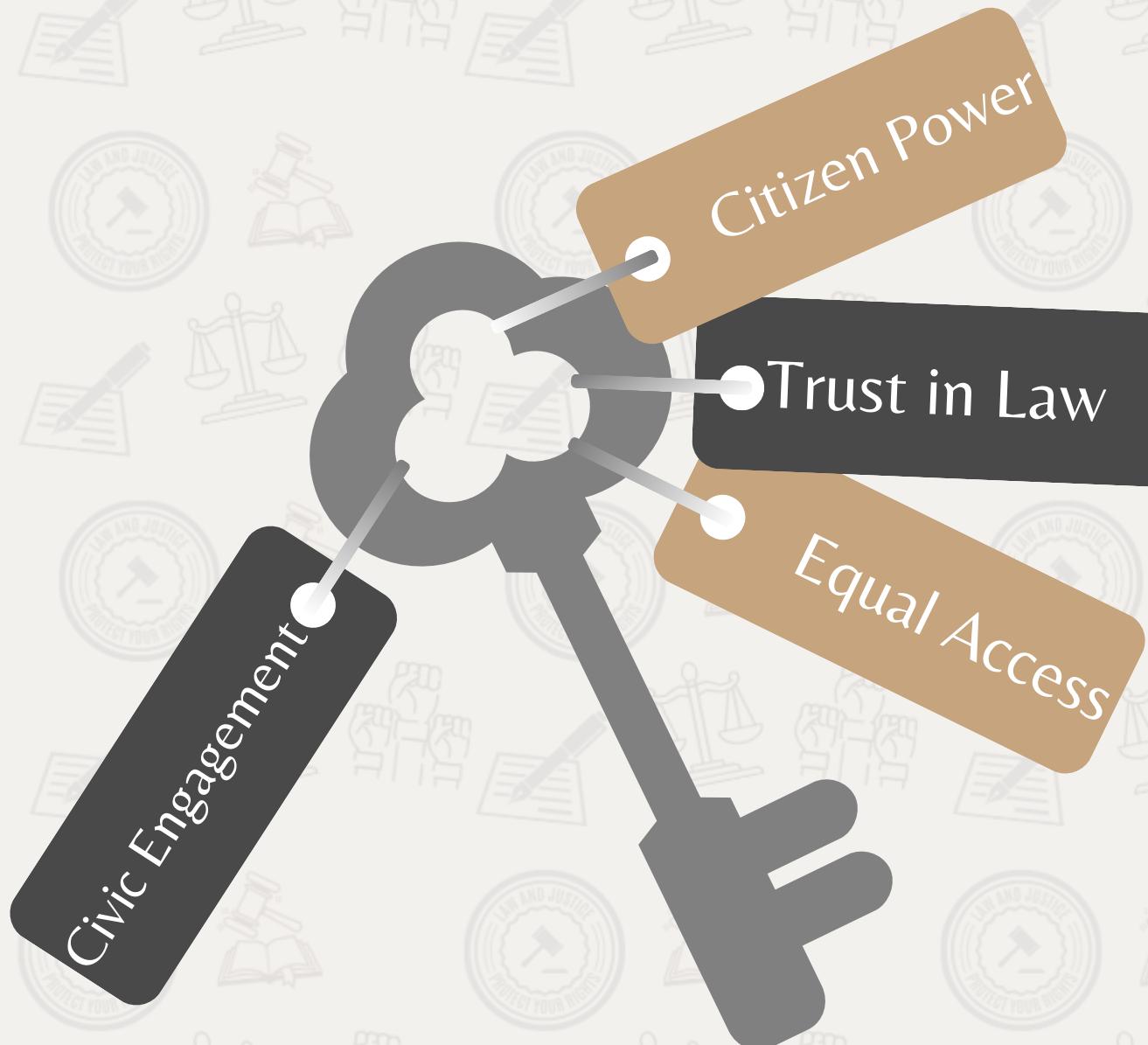
To achieve optimal performance, the chosen LLM will be fine-tuned on a massive corpus of legal documents from various Indian courts. This focused training equips the model to grasp the intricacies of legal jargon and reasoning, leading to more accurate and legally relevant summaries.

- **Iterative Refinement:**

The system will undergo continuous improvement through a feedback loop. User feedback regarding summary quality and relevance will be incorporated to further refine the model's performance. Additionally, the system can be trained on newly available legal judgments, ensuring it stays up-to-date with evolving legal language and trends.

SOCIETAL IMPACT

This legal judgment summarization system reaches far beyond simply understanding legalese. It has the potential to create a significant positive ripple effect across society in several key ways:



1) Democratization of Legal Knowledge

This system dismantles the legal knowledge gap, empowering citizens with clear summaries in local languages. This fosters a more informed and active citizenry, ready to participate in legal processes.

2) Enhanced Civic Engagement

Legal knowledge fuels civic engagement! Understanding legal concepts empowers citizens to participate in discussions, hold officials accountable, and advocate for positive legal reforms.

3) Strengthened Community Confidence

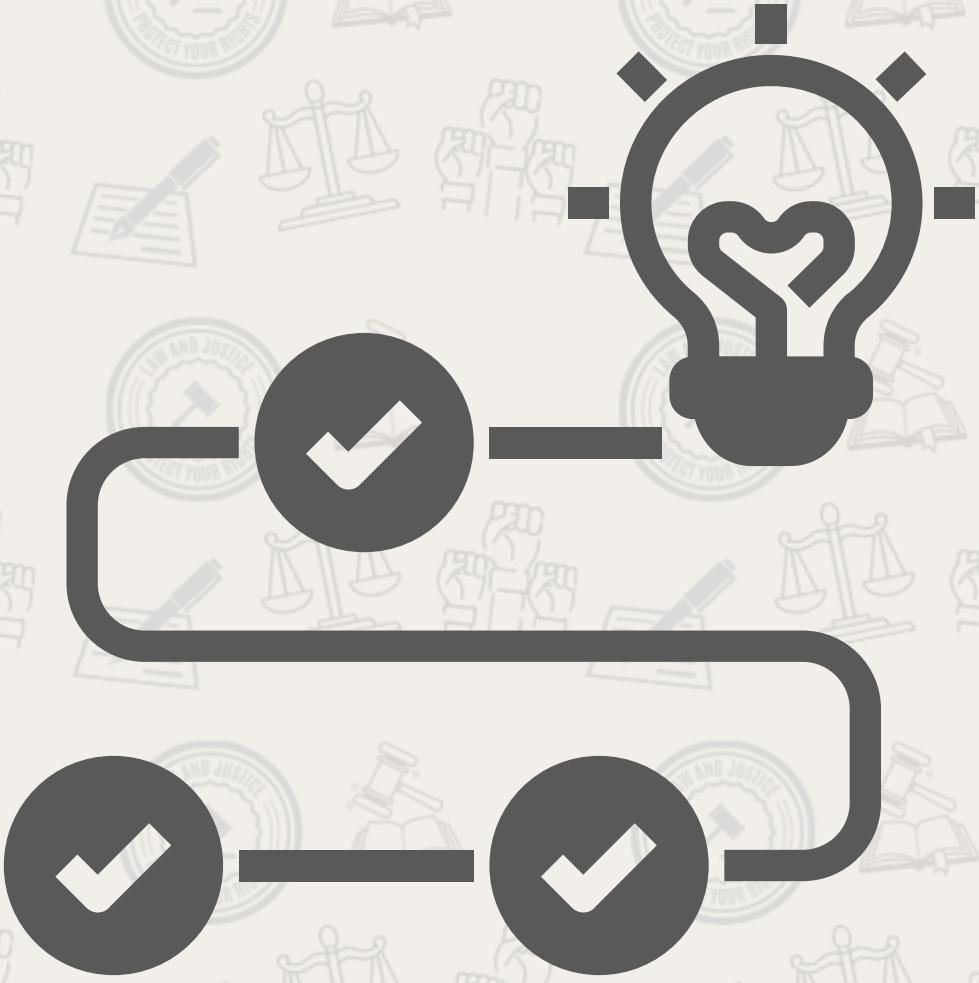
Clear legal summaries build trust! Citizens empowered to understand court proceedings can hold institutions accountable and advocate for fair legal practices in their communities.

4) Reduced Socioeconomic Disparities:

Legal complexities become equal footing! Clear summaries bridge the gap for low-income individuals, promoting equal access to justice without relying on expensive legal counsel.

CONCLUSION

In conclusion, this proposed legal judgment summarization system leverages artificial intelligence to bridge the gap between complex legal language and the everyday citizen. By generating clear, concise summaries in diverse local languages, the system empowers individuals to understand their legal rights and navigate the legal system with greater confidence. This not only promotes individual empowerment but fosters a more inclusive and transparent legal landscape. We believe this project has the potential to significantly enhance public access to justice and contribute to a more informed and engaged citizenry.



THANK YOU