

RAJASTHAN POLICE HACKATHON 1.0



Analysis of **FIR** using **AI/ML** for proper Act and Section

Team: NeuralNinjas

Reg. ID: 765

Problem Statement-3

GitHub Repo: [Here](#)





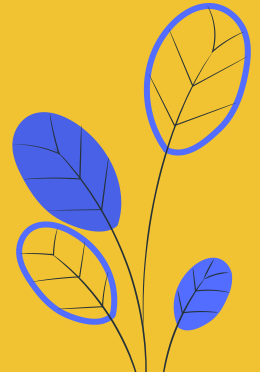
PROBLEM STATEMENT

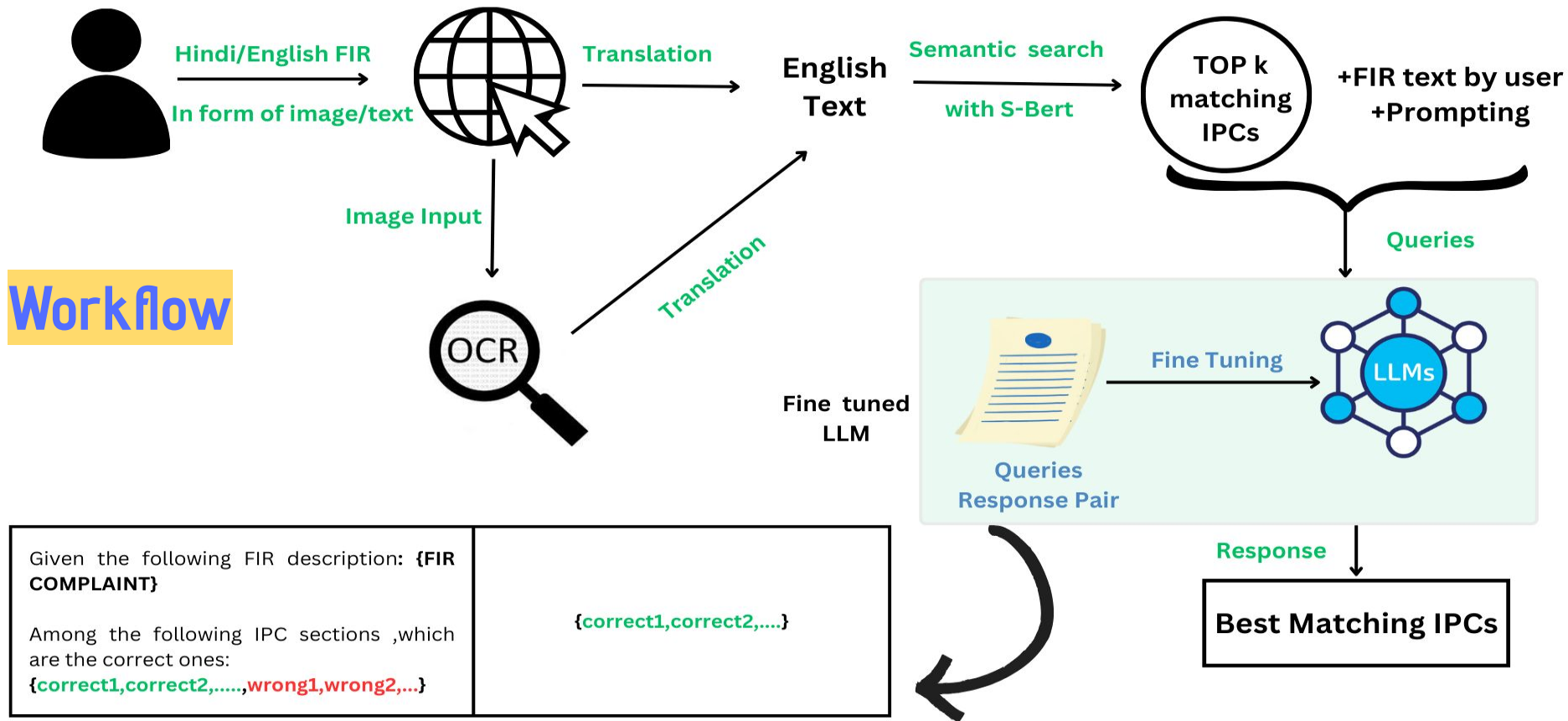
Develop an AI-powered solution to address the precise identification of legal sections within First Information Reports (FIRs)





OUR SOLUTION : NyayaFinder





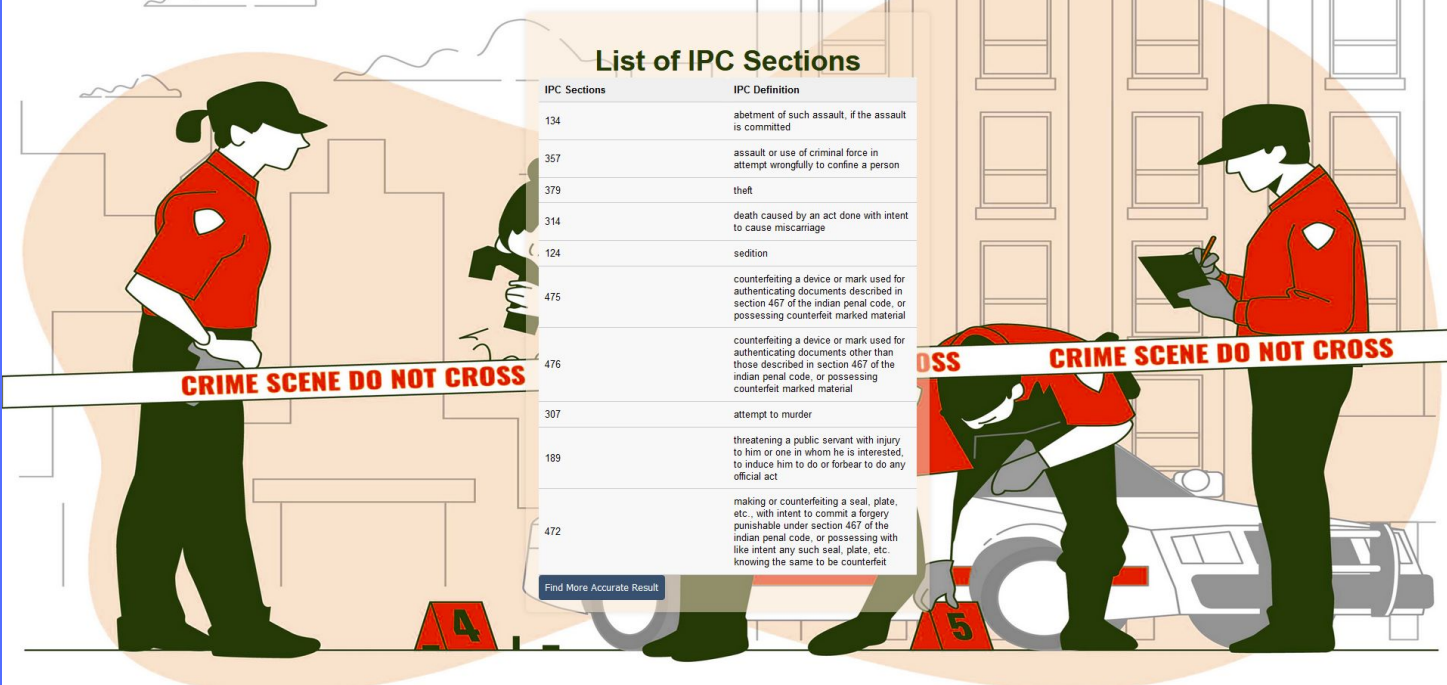
PRODUCT DEMO

1. User gives Input in form of text or image
Languages supported: Hindi, English



WEBSITE DEMO

2. Top K IPCs found by semantic search



The illustration depicts a crime scene investigation. Two police officers in red uniforms and black caps are present. One officer stands on the left, looking towards the center. The other officer stands on the right, holding a clipboard and pen, looking down at it. A white police car is parked on the right side. A red and white striped barrier with the text "CRIME SCENE DO NOT CROSS" is stretched across the scene. In the center, a table titled "List of IPC Sections" is displayed. The table has two columns: "IPC Sections" and "IPC Definition". The table lists 11 IPC sections with their corresponding definitions. A red evidence marker with the number "4" is on the ground near the left officer, and a red evidence marker with the number "5" is on the ground near the right officer. A blue button with the text "Find More Accurate Result" is located at the bottom of the table.

IPC Sections	IPC Definition
134	abetment of such assault, if the assault is committed
357	assault or use of criminal force in attempt wrongfully to confine a person
379	theft
314	death caused by an act done with intent to cause miscarriage
124	sedition
475	counterfeiting a device or mark used for authenticating documents described in section 467 of the indian penal code, or possessing counterfeit marked material
476	counterfeiting a device or mark used for authenticating documents other than those described in section 467 of the indian penal code, or possessing counterfeit marked material
307	attempt to murder
189	threatening a public servant with injury to him or one in whom he is interested, to induce him to do or forbear to do any official act
472	making or counterfeiting a seal, plate, etc., with intent to commit a forgery punishable under section 467 of the indian penal code, or possessing with like intent any such seal, plate, etc. knowing the same to be counterfeit

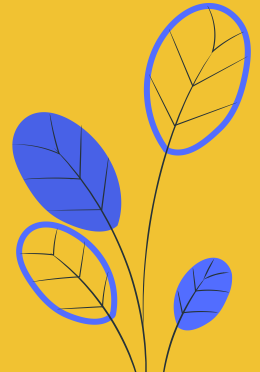
[Find More Accurate Result](#)

PRODUCT DEMO

3. Best Matching IPCs found by fine-tuned LLM (Final Output)



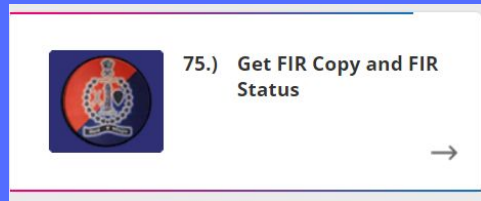
TECHNICALITIES



Data Collection

Jan Soochna Portal

<https://iansoochna.rajasthan.gov.in/Services/DynamicControls>



FIR PDFs written in Hindi

OCR
(Tesseract)

Translation

(Google Translator API
Free- dividing into chunks)

1. IPC Sections Dataset

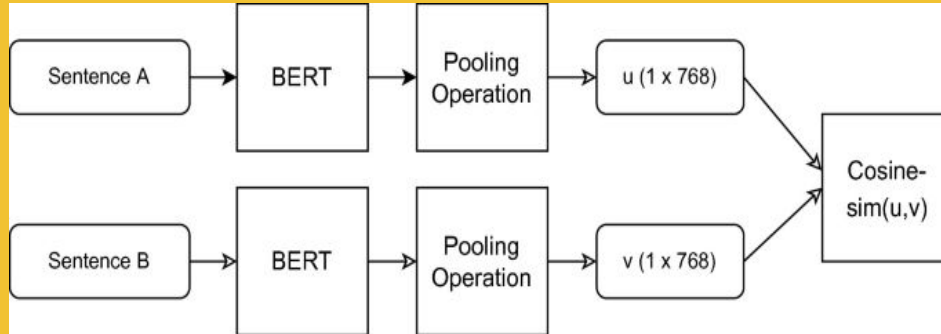
(Provided By Mentor)

2. FIR Complaint Dataset
in English



LLM comes into Picture

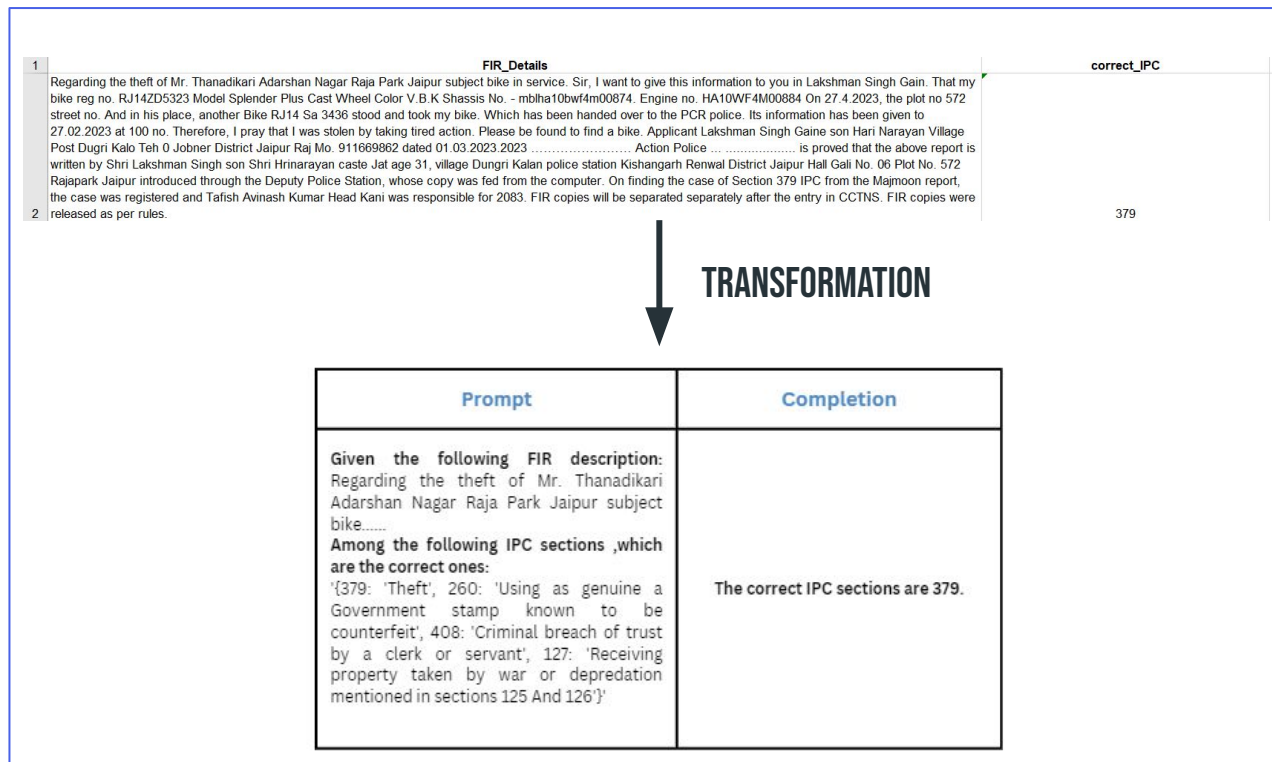
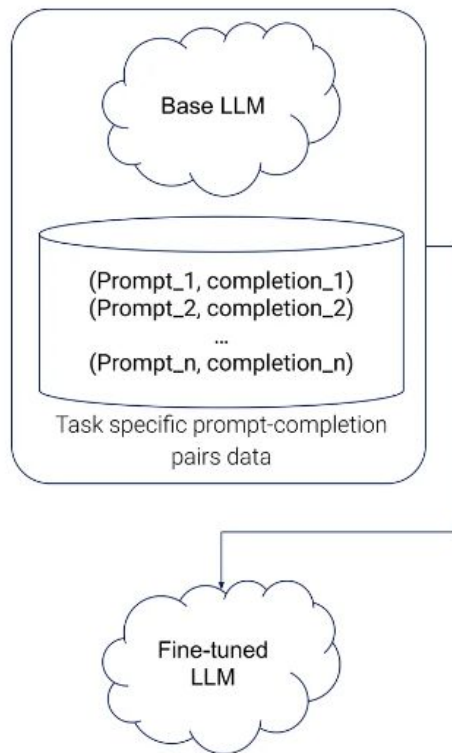
SBERT is considered the best choice for measuring sentence similarity because it combines the benefits of contextual understanding, semantic representation, versatility, and strong performance on various NLP tasks.



Using LLMs like LLama provides superior results compared to conventional similarity search methods because LLMs have the ability to understand the semantic meaning, context, and intricacies of legal text in FIR complaints.



Fine Tuning Llama 2



A DATASET OF 300+ SUCH PROMPT,COMPLETION PAIRS ARE USED TO FINE TUNE LLAMA 2

What is the need of S-BERT ?

- Llama is constrained by a token limit of 2048, making it impractical for comparing all IPC sections in one go.
- We integrate S-BERT to identify the top semantically matching IPC sections, optimizing Llama's token allocation.



Best Matching IPC with Fine-tuned LLM

IPC Sections	IPC Definition
134	abetment of such assault, if the assault is committed
357	assault or use of criminal force in attempt wrongfully to confine a person
379	theft
314	death caused by an act done with intent to cause miscarriage
124	sedition
475	counterfeiting a device or mark used for authenticating documents described in section 467 of the indian penal code, or possessing counterfeit marked material
476	counterfeiting a device or mark used for authenticating documents other than those described in section 467 of the indian penal code, or possessing counterfeit marked material
307	attempt to murder
189	threatening a public servant with injury to him or one in whom he is interested, to induce him to do or forbear to do any official act
472	making or counterfeiting a seal, plate, etc., with intent to commit a forgery punishable under section 467 of the indian penal code, or possessing with like intent any such seal, plate, etc. knowing the same to be counterfeit

Top K iPCs found

FINE-TUNED LLM

Final Result

The correct IPC sections are 124, 134, 307, 314, 357, 379.

Final match using LLM

Key Success Factors of Our Effective Solution



01 Scalability

Our product efficiently handles long FIR complaints by dividing them into manageable chunks

02 Accuracy

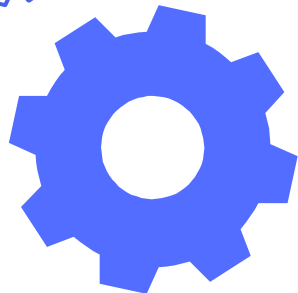
When the embeddings include most relevant IPCs, our LLM effectively identifies and matches the correct IPCs with high precision and accuracy

03 Processing Speed

lightning-fast processing speed for embedding, and when coupled with LLM, it completes the task in less than 1 minute, offering both speed and efficiency.

04 Cost Effectiveness

Our solution is offered completely free, ensuring accessibility to all users without any cost constraints.



THANKS

We are open to questions.

