RAJASTHAN POLICE HACKATHON 1.0

Basic details of the Problem Statement and Team:

Problem Statement: Analysis of FIR using AI/ML for proper Act and Section.
(*Advanced FIR Processing System)*

Team Name: Future coders

Team Members are: Mohit Mittal
Anima Johri
Akshita Mishra
Lucky Panchal

CHALLENGES FACED

- CCTNS Datasets Not Utilized: ML hasn't been applied to CCTNS datasets, limiting the use of FIRs available on Raj-cop.

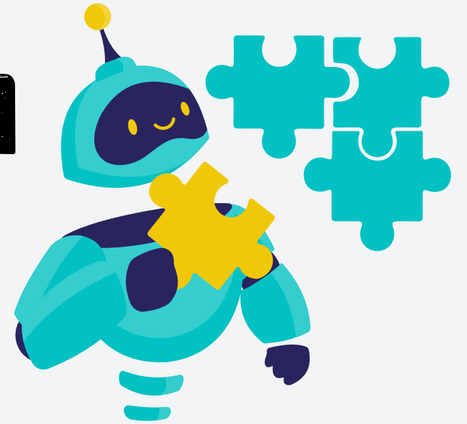
- Local and Regional Language Datasets: FIR datasets are offered in local and regional languages, presenting a challenge for processing.

CHALLENGES FACED

- Traditional document processing methods often fall short when it comes to handling the increasing volume and complexity of both physical and digital paperwork

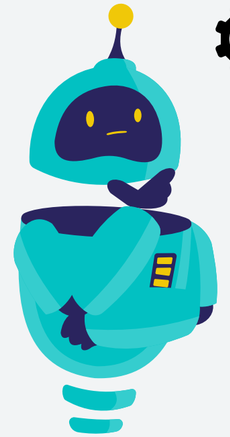
- Organizing information into relevant sections poses challenges in FIRs

OBJECTIVES OF THE AI/ML SYSTEM



- **Natural Language Processing (NLP):** Develop NLP algorithms to comprehend and extract crucial details from complainant-provided information.
- **Structured Data Representation:** Transform extracted information into a structured format for consistent and clear FIR generation.
- **Accuracy and Completeness:** Ensure the AI/ML system accurately captures all necessary information, minimizing errors and omissions in FIRs.
- **Multi-language Support:** Implement language-agnostic capabilities to support FIR drafting in multiple languages used within the jurisdiction.
- **Legal and Regulatory Compliance:** Design the system to adhere to important legal requirements and standard FIR formats specified by law enforcement agencies.

IDEA/APPROACH DETAILS

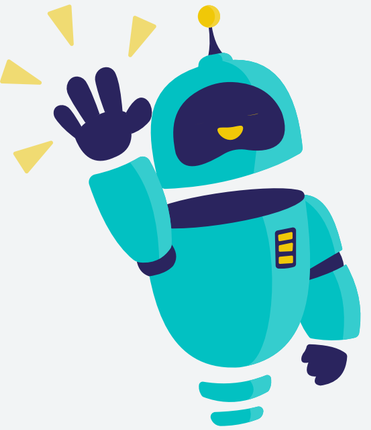


1. Handling Local and Regional Language Datasets:

Overcome the language barrier by implementing NLP techniques to process FIR datasets available in local and regional languages.

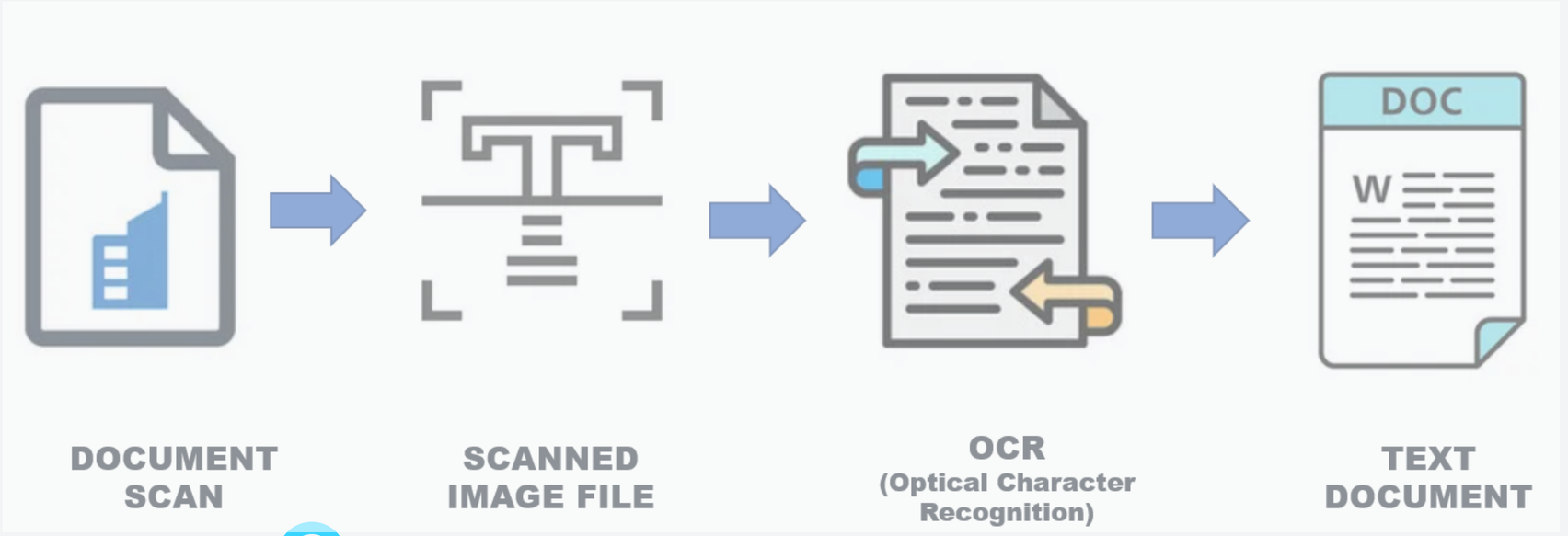
- Implement NLP algorithms for language translation to convert local and regional languages to a standardized language (e.g., English).
- Develop language-specific models to process and understand the context of FIRs in various languages.
- Create a multilingual NLP model that can comprehend and extract information from FIRs in diverse languages.

Digitization of FIR Documents:



Overcome the language barrier by implementing NLP techniques to process FIR datasets available in local and regional languages.

- Run the OCR process to extract text from the FIR documents. The OCR software will analyze the images and convert the text into a machine-readable format.



Division by sections

Build an integrated system that combines ML for CCTNS datasets and NLP for local language datasets to provide a comprehensive solution.

- Train a machine learning model to classify the FIR into the appropriate legal Act and Section based on the extracted text.
- Use historical FIR data to train the model, allowing it to recognize patterns and correlations between text and legal classifications.
- Extract relevant features from the FIR text, such as keywords, entities, and context, to enhance the model's ability to accurately classify Acts and Sections.

STEPS TO DEVELOP AN ADVANCED FIR PROCESSING SYSTEM

Legal Expertise

Research and Data Collection

Natural Language Processing (NLP)

Machine learning models

Legal dataset Integration

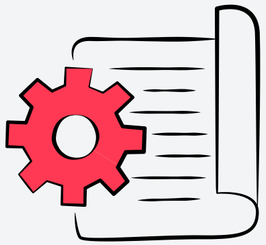
User Interface Design

Testing and Validations

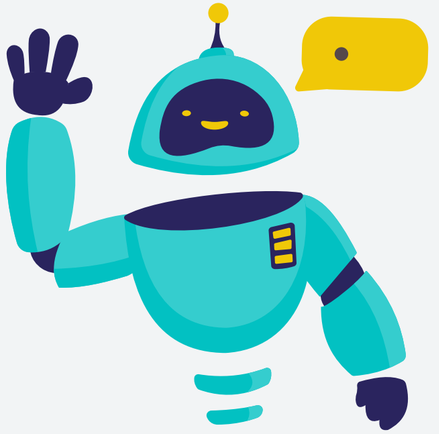
Security and compliance

Collaboration with Authorities

Continuous Improvement



TECHNOLOGY STACK



Programming Languages: Python, Java, C++

Frameworks: TensorFlow, PyTorch

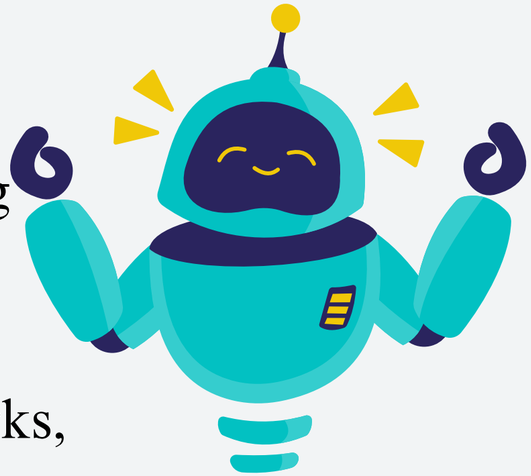
Tools: Open CV for computer vision

Speech synthesis libraries: eSpeak, google Text-to-speech(gTTS)

Cloud Sevices: AWS, Google cloud platform

Mobile App Development (if appicable): Android Studio, Swift

FUTURE ASPECTS



- Extend the system to support global legal frameworks by incorporating models and databases that can handle diverse legal systems and languages.
- Explore the application of deep learning models, such as neural networks, to further enhance the accuracy of Act and Section classification.
- Incorporate multimodal analysis by integrating image and audio processing capabilities.
- Work towards achieving real-time processing capabilities to enable law enforcement agencies to swiftly analyze and respond to incidents.