RAJASTHAN POLICE

LEGAL AI
ASSISTANT

Problem Statement

Design an AI/ML system with OCR capabilities, state-specific legal rules integration, and multi-language support to automate and enhance the accuracy of FIR drafting, addressing challenges of underutilized CCTNS datasets, diverse regional languages, and legal section multiplicity, with a focus on transparency, accountability, and compatibility with existing systems.

- AI - IPC AND CRPC SECTION IDENTIFIER

- MULTI-LANGUAGE SUPPORT

- EASY UI

- BILINGUAL OCR CAPABILITY

- TRI-WAY INPUT METHOD

SOLUTION

- Created webpages using HTML, CSS, and Django framework.
- Developed a Streamlit UI with models for IPC and CRPC section identification using NLP and LLM.
- Implemented OCR for English and Hindi printed and handwritten text documents.
- Incorporated a voice text input model for diverse input methods.
- Stored data in MongoDB database in XML format for structured information.
- Established a dashboard using the open-source Streamlit platform.

The IPC,CRPC Section Identifier

- **IPC Section Identifier:**
  - A model trained to predict the specific section of the Indian Penal Code (IPC) based on textual descriptions of legal incidents in JSON format.
  - Utilizes a neural network trained on labeled IPC section data for accurate predictions.

- **CRPC Section Identifier:**
  - Similar to the IPC Section Identifier, this model predicts the section of the Code of Criminal Procedure (CRPC) based on textual descriptions in JSON format.
  - Leverages a neural network trained on labeled CRPC section data, ensuring precise section identification.
  -

Predicted IPC Section Information:

	426
chapter	16
chapter_title	sexual offences
Section	376AB
section_title	Punishment for rape on woman under twelve years of age
section_desc	Whoever, commits rape on a woman under twelve years of .
label	510

Predicted CRPC Section Information:

	460
chapter	32
section	424
section_title	Suspension of execution of sentence of imprisonment
section_desc	When an offender has been sentenced to fine only and to imprisonment in default of
label	404

OCR (Printed and Handwritten)

- Our solution incorporates Optical Character Recognition (OCR), a transformative technology that converts scanned paper documents, PDFs, and images into editable and searchable data. This versatile OCR functionality extends to recognizing both printed and handwritten text, enabling seamless extraction of information from diverse document types. While specific code snippets are not provided, OCR is a crucial component empowering the system to process and extract details efficiently. With applications across various use cases, including legal incidents, OCR enhances data accessibility from non-digital formats.

Voice Text Input

- This refers to the capability to convert spoken words into text. Users can provide input to the system using their voice, and the system processes and analyzes the spoken words to generate corresponding text data.

Performance Metrics

IPC SECTION IDENTIFIER

Accuracy : 97.39

CRPC SECTION IDENTIFIER
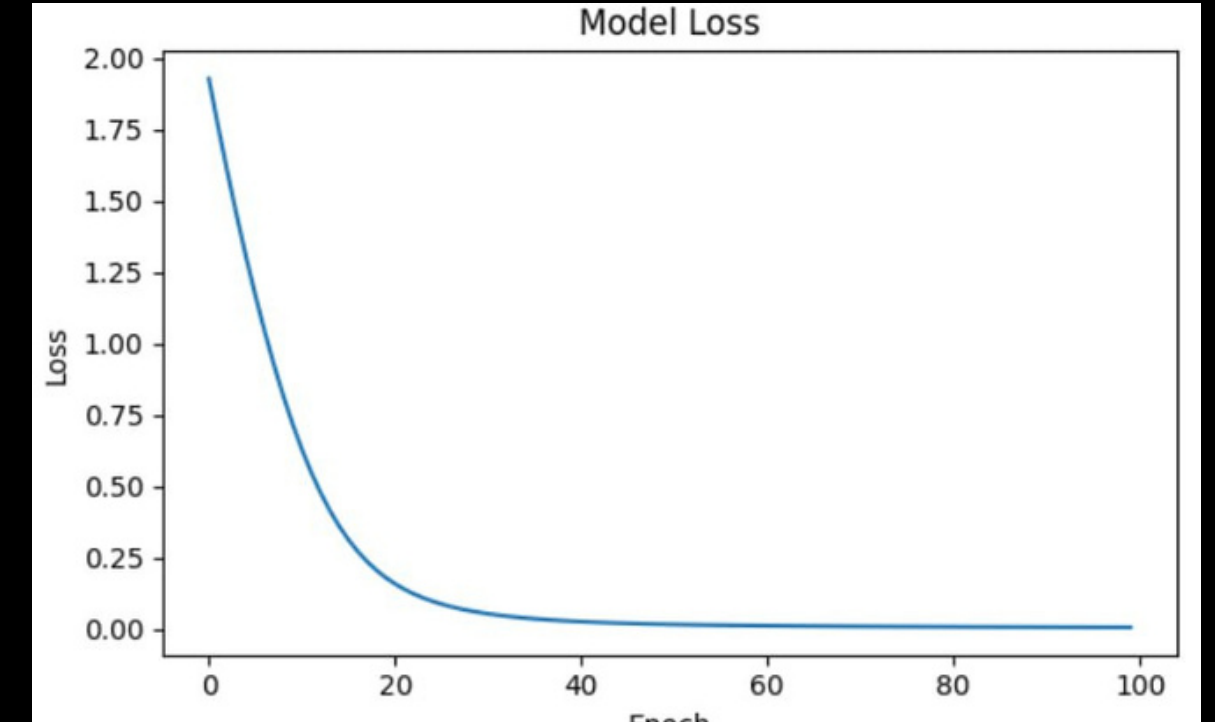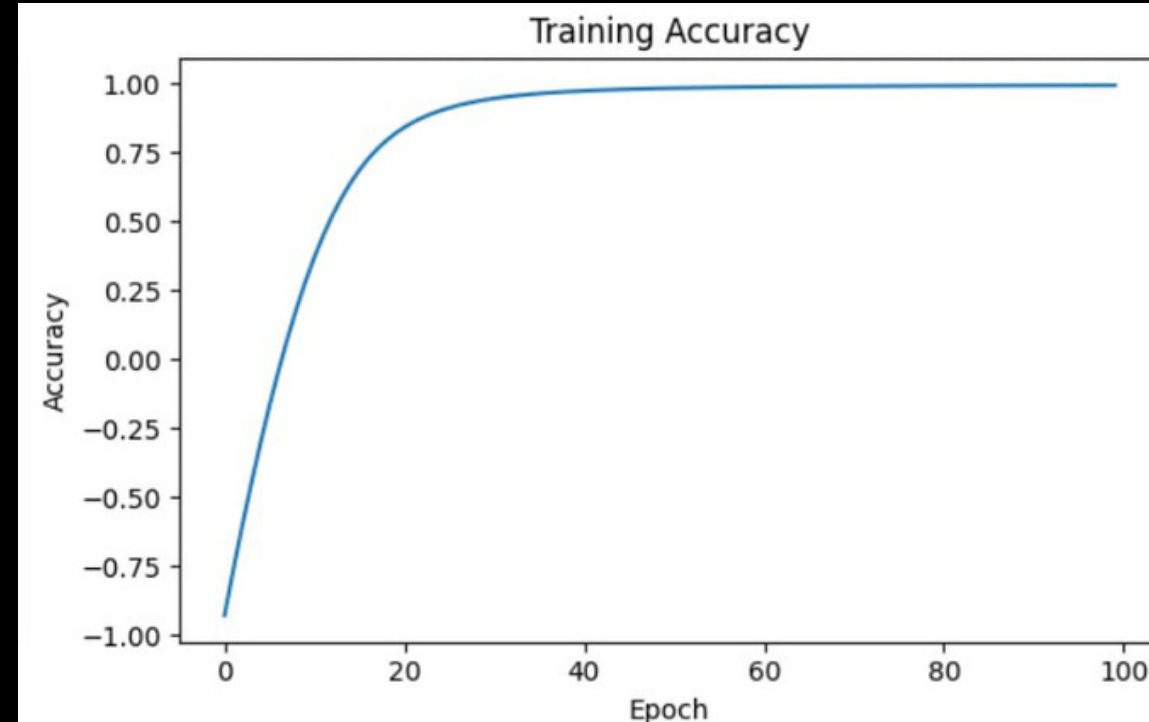
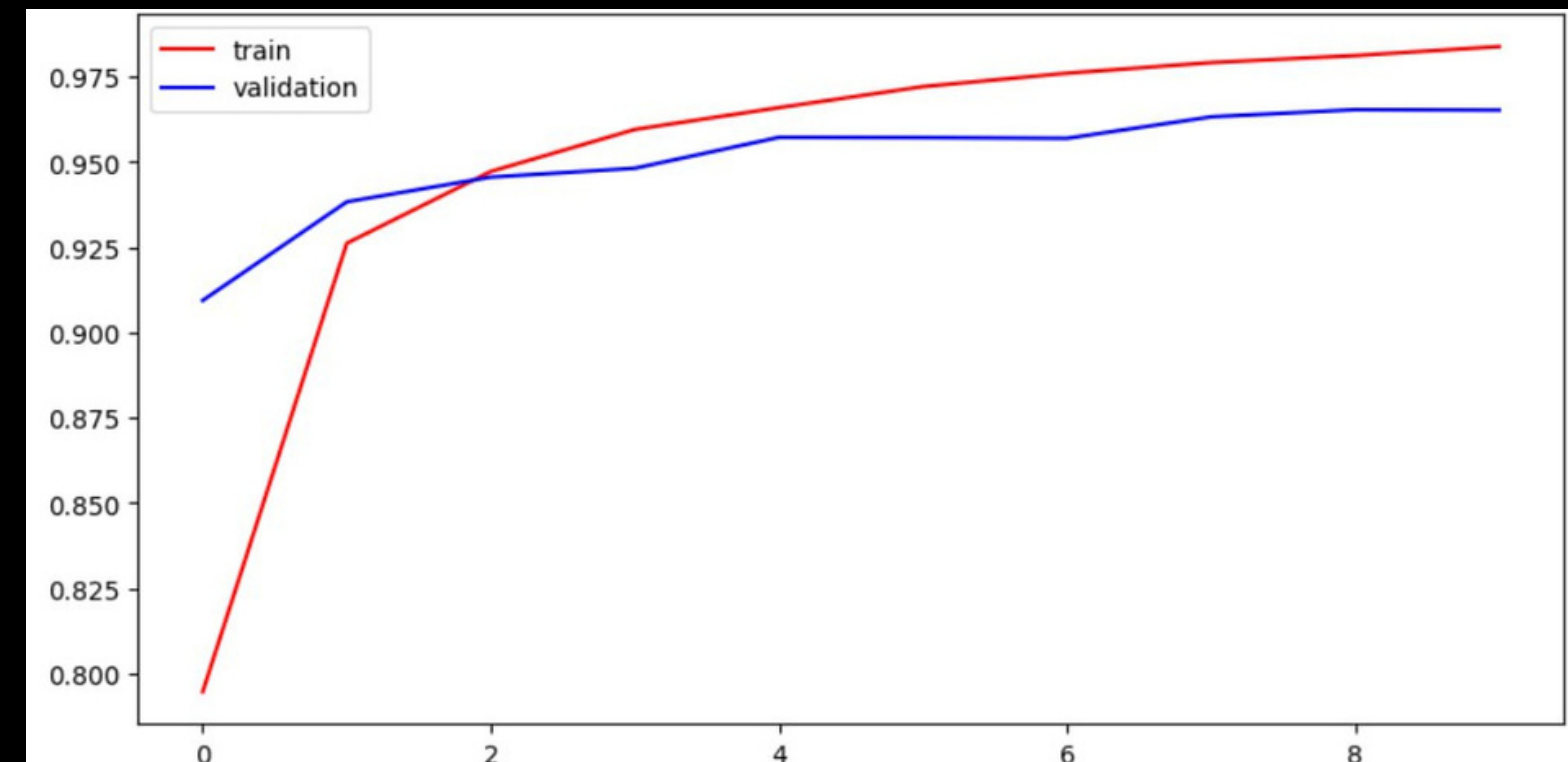Accuracy : 99.76

File Upload OCR

Accuracy : 95.42

Voice Recognition

Accuracy : 99.46

Graphical Representation

Voice Recognition ⟶



Optical Character Recognition ⟶

Streamlit UI Integration in Solution

Streamlit, a Key Player in User Interaction
Our solution leverages the power of Streamlit to provide a seamless and interactive user interface, enhancing the overall user experience in FIR drafting.
It provides the following:

- **User-Friendly Interface**
- **Real-Time Updates**
- **Multi-Model Integration**
- **Flexibility for IOs**
- **Accessibility and Compatibility**
- **Continuous Improvement**
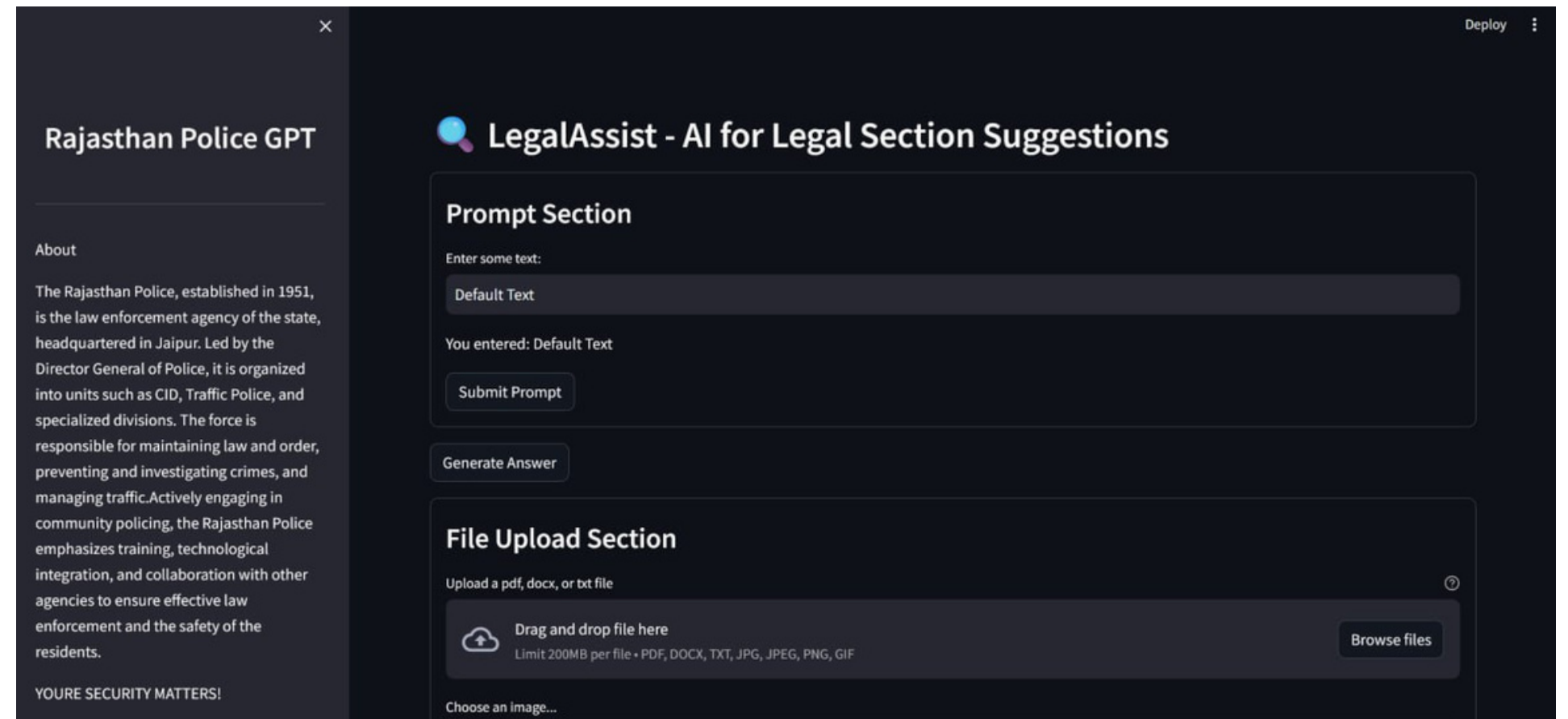
POTENTIAL OUTCOMES

Best Case Scenario

- Provides Seamless Integration
- High Accuracy
- Easy User Adoption
- Time and Resource Savings

Realistic Success Scenario

- Effective Section Identification
- User Acceptance
- Streamlined Process
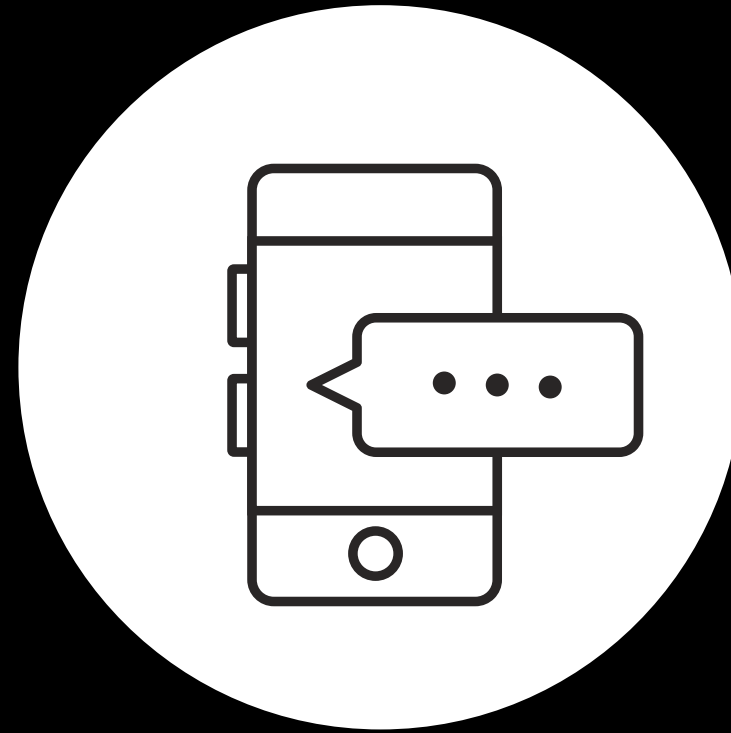- Improved Data Analysis

Worst Case Scenario

- Integration Challenges
- Model Accuracy Issues
- Limited User Adoption
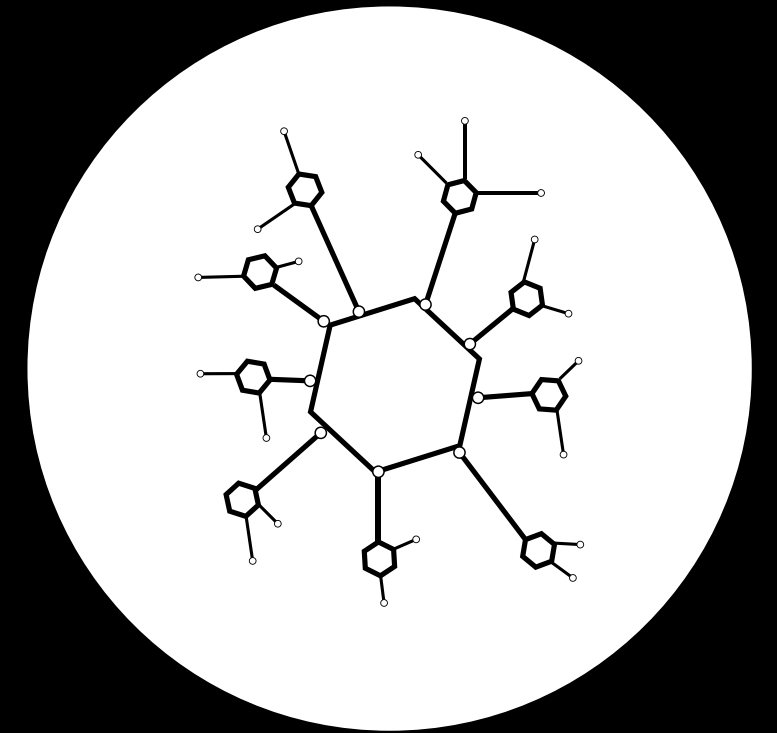- Operational Disruptions

FUTURE IDEAS

SEMANTIC REPRESENTATION

CHATBOT ANALYZER

BLOCKCHAIN SUPPORT

Dashboard :

URL - https://app.powerbi.com/reportEmbed?reportId=12f153df-e3b3-4b39-8b59-2b6bb5d87513&autoAuth=true&ctid=3e19468b-e065-44a1-9cde-749457855c06

THANK YOU

Average of Number of Incidents by Crime Type and Number of Incidents

25.00 (6.67%)
30.00 (8%)
40.00 (10.6...)
150.... (40%)

Crime Type
- Theft
- Burglary
- Assault
- Vandalism
- Robbery
- Fraud

50.00 (13.33%)
80.00 (21.33%)
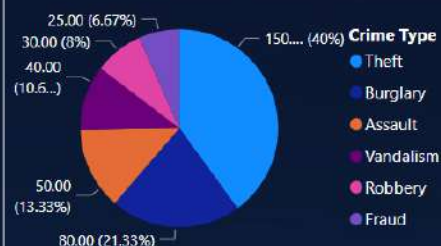
Number of Incidents by Year

Number of Incidents

100

50

0

2016   2018   2020

Year
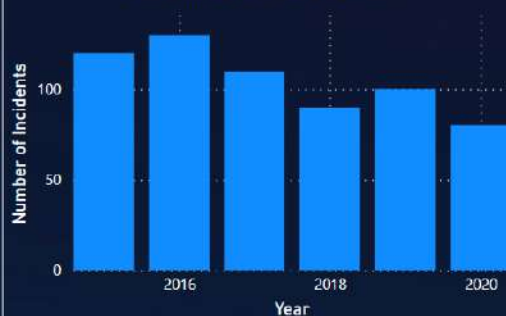
Number of Cases by Case Status

50 (13.16...)

Case Status
- Solved
- Pending
- Closed
- Dismissed

100 (26.32%)
200 (52.63%)

Sum of Female Incidents and Sum of Male Incidents by Age Group

● Sum of Female Incidents   ● Sum of Male Incidents

Age Group

18-25
26-35
36-45
46-55

0   50   100   150

Sum of Female Incidents and Sum of Male Incidents

LEGAL ASSITANT

Primary Actor

complaintee

Secondary Actor

Centralized Database

Login / Register

<<include>>

Enter details

<<include>>

Contact details

Id no

Emails

Ph no

Details of the user

Complaint

<<include>>

Details from the complaitee

District

Description

Contact details

LLM - RAG MODEL

IPC/CRPC IDENTIFIER

Voice identifier for recommendation of ipc/ crpc

Document ocr recongition

Secondary Actor

Centralized Database

PRIMARY ACTOR - initiate use cases and interact with the system
SECONDARY ACTOR - assists or facilitates the primary actor
INCLUDE - extend the base use case
EXCLUDE - optional use case which extends the base use case