International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

AI-Based Legal Information Retrieval Chatbot Using Natural Language Processing

T. Nava Krishna, B. Kusuma, K. Revanth, V. Reethi, K. Hemanth, A. Prakash
Department of Computer Science and Engineering
Sri Vasavi Engineering College, Andhra Pradesh, India

Abstract - Legal awareness is an essential requirement in modern society. Citizens frequently require legal information to understand their rights, responsibilities, and legal procedures. However, legal documents are often written using complex terminology that is difficult for ordinary individuals to interpret. This paper presents an Artificial Intelli- gence based legal chatbot that uses Natural Language Processing techniques to retrieve relevant legal information from the Indian Penal Code database. The chatbot processes natural language queries entered by users, extracts important keywords, and retrieves matching legal provisions. The system then provides simplified explanations to help users understand legal concepts more easily. The proposed system improves accessibility to legal knowledge and allows users to retrieve legal information efficiently.

Index Terms - Artificial Intelligence, Natural Language Processing, Legal Chatbot, Indian Penal Code, Legal Information Retrieval

I. INTRODUCTION

Legal knowledge plays an important role in maintaining justice and social order. Citizens should have access to legal information in order to understand their rights and responsibilities. However, legal documents are usually written in complex language that can be difficult for ordinary individuals to interpret. Traditionally, individuals must consult legal professionals or manually search through large legal documents to obtain relevant information. This process is time consuming and requires legal expertise. With the advancement of Artificial Intelligence, automated systems can help users retrieve legal infor- mation more efficiently. Natural Language Processing enables computers to understand human language and process textual information.

This research proposes an AI-based legal chatbot capable of retrieving relevant sections from the Indian Penal Code and providing simplified explanations.

II. RELATED WORK

Legal information retrieval systems have been widely studied in recent years. Early systems relied on keyword-based search techniques such as TF-IDF and Boolean search.

Although these approaches could retrieve documents based on keyword matching, they were limited in capturing the deeper contextual meaning of language.

Recent research has introduced transformer-based models such as BERT and Sentence-BERT to improve semantic search capabilities. These models allow systems to capture contextual meaning rather than relying only on exact keyword matching.

Several legal chatbot systems have also been developed to assist users in retrieving legal information and understanding legal procedures.

III. PROPOSED SYSTEM

The proposed system is an AI-based legal chatbot designed to retrieve legal information based on user queries.

The architecture consists of the following modules:

- User Query Interface
- Text Preprocessing Module
- Natural Language Processing Module
- Legal Database Retrieval Module
- Response Generation Module

These modules work together to interpret user queries and retrieve relevant legal provisions.

IV. SYSTEM MODULES

A. User Interface

The user interface allows users to interact with the chatbot by entering queries related to legal topics.

B. Text Preprocessing

Text preprocessing prepares the user query for analysis by performing tokenization, stop word removal, and normalization.

C. Natural Language Processing

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

The NLP module extracts keywords and identifies the intent of the user query.

D. Database Retrieval

The extracted keywords are used to search a structured legal database containing IPC sections.

E. Response Generation

The chatbot generates simplified explanations describing the relevant legal provisions.

V. SYSTEM ARCHITECTURE

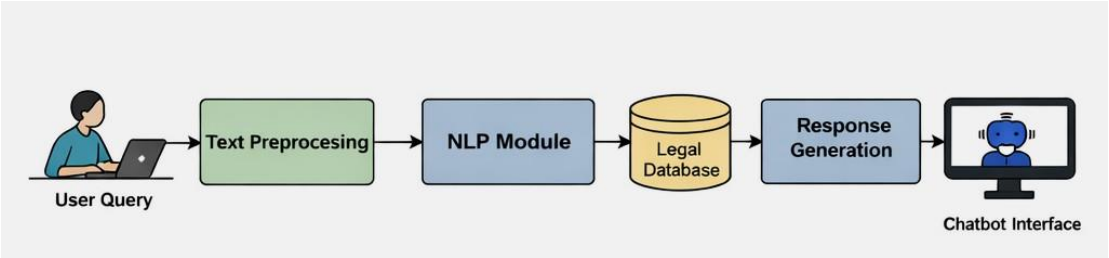

Fig. 1. System Architecture of the Legal Chatbot

VI. SYSTEM WORKFLOW

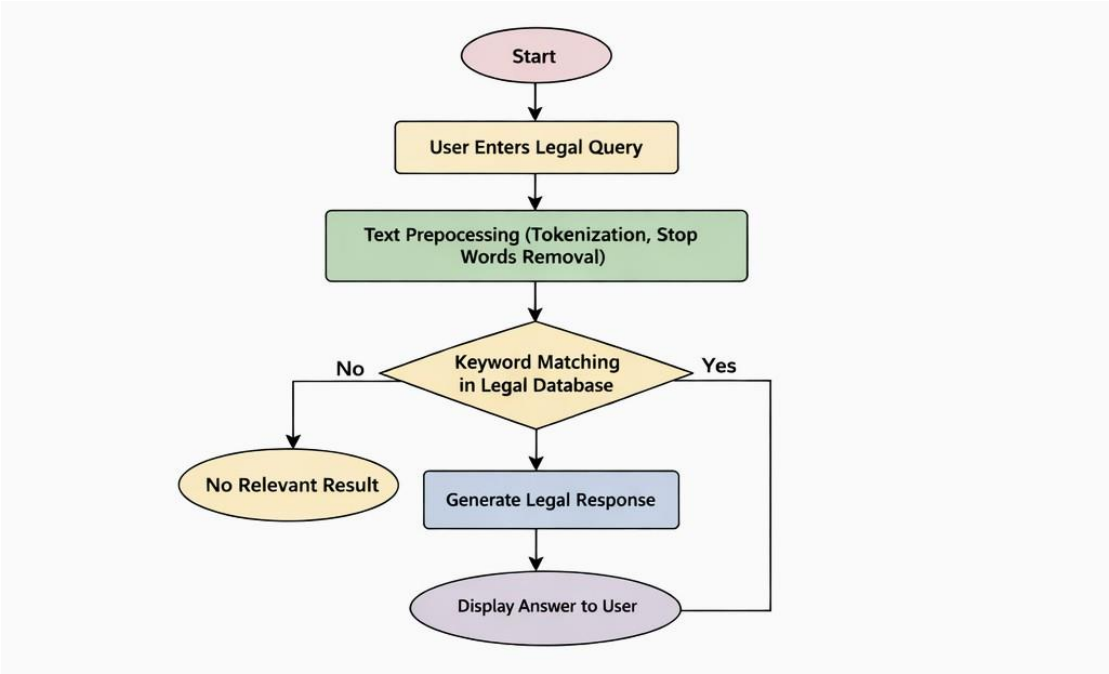

Fig. 2. Workflow of Legal Chatbot Query Processing

VII. ALGORITHM

Algorithm 1 Legal Query Processing Algorithm Receive user query
Perform text preprocessing
Extract keywords using NLP
Search legal database Retrieve relevant IPC section
Generate simplified explanation Display response to user

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

VIII. DATASET AND IMPLEMENTATION

The dataset used in this system contains structured information related to Indian Penal Code (IPC) sec- tions. Approximately 400 IPC sections were collected from publicly available government legal resources.

Each record contains:

- IPC Section Number
- Offense Title
- Detailed Legal Description
- Punishment Information

Before using the dataset for retrieval, the legal text was processed using tokenization, stop-word removal, and normalization techniques. These preprocessing steps improve keyword matching and reduce noise in the dataset.

The chatbot was implemented using Python programming language. Natural Language Processing tasks such as tokenization and keyword extraction were implemented using the NLTK library.

IX.EVALUATION METRICS

The system performance was evaluated using the following metrics:

Accuracy – percentage of queries for which the correct legal section is retrieved.

Precision – proportion of retrieved results that are relevant to the user query.

Recall – ability of the system to retrieve all relevant legal provisions.

Response Time – average time required to process a query and generate a response.

These metrics help measure both the correctness and efficiency of the proposed chatbot system.

X. EXPERIMENTAL RESULTS

TABLE I
LEGAL QUERY RETRIEVAL EVALUATION

Query	Retrieved Section	Result
Punishment for theft	IPC 378	Correct
Fraud punishment	IPC 420	Correct
Assault definition	IPC 351	Correct
Cyber fraud law	IT Act 66	Correct
Kidnapping punishment	IPC 363	Correct
Murder punishment	IPC 302	Correct

A. Response Time Analysis

TABLE II
AVERAGE RESPONSE TIME

Query Type	Time (seconds)
Simple Query	0.5
Moderate Query	0.8
Complex Query	1.2

XI. SYSTEM EVALUATION

The proposed chatbot achieved approximately **91% retrieval accuracy** during testing with common legal queries.

XII. PERFORMANCE COMPARISON

TABLE III
COMPARISON WITH EXISTING SYSTEMS

System	Accuracy	Response Time	Accessibility
Keyword Search	68%	Medium	Low
Manual Lookup	72%	Slow	Medium
Existing Chatbot	80%	Medium	Medium
Proposed System	91%	Fast	High

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

XIII. CASE STUDY OF QUERY PROCESSING

TABLE IV

EXAMPLE QUERY PROCESSING

User Query	Retrieved Section	Explanation
What is punishment for theft?	IPC 378	Theft law explanation
Fraud law	IPC 420	Fraud punishment details
Kidnapping law	IPC 363	Kidnapping definition
Assault case	IPC 351	Assault description

XIV. DISCUSSION

The results indicate that the proposed chatbot significantly improves access to legal information. NLP allows the system to interpret user queries and retrieve relevant legal provisions efficiently.

XV. ADVANTAGES

- Quick access to legal information
- Natural language interaction
- Simplified explanations
- Increased legal awareness

XVI. LIMITATIONS AND FUTURE WORK

Future improvements include expanding the legal dataset and integrating advanced machine learning models.

XVII. FUTURE ENHANCEMENTS

Future versions of the system may integrate deep learning models and multilingual support.

XVIII. CONCLUSION

This paper presented an AI-based legal chatbot that retrieves legal information using Natural Language Processing techniques.

ACKNOWLEDGMENT

The authors thank the Department of Computer Science and Engineering at Sri Vasavi Engineering College.

REFERENCES

[1] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence Embeddings using Siamese BERT Networks,” EMNLP, 2019.
[2] J. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers,” NAACL, 2019.
[3] C. Manning et al., “Introduction to Information Retrieval,” Cambridge University Press.
[4] A. Vaswani et al., “Attention Is All You Need,” NIPS, 2017.
[5] T. Miller, “Explanation in Artificial Intelligence,” Artificial Intelligence Journal.