IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A BERT-Based Intelligent Chatbot Framework For Judicial Assistance

Dr. Bama S
Professor, Dept of CSE
Sri Venkateshwara College of Engineering. Bengaluru, India

Mr Abhishek V M
Student, Dept of CSE
Sri Venkateshwara College of Engineering. Bengaluru, India

Mr Abhi Puttaswamy M P
Student, Dept of CSE
Sri Venkateshwara College of Engineering. Bengaluru, India

Mr Hruday V B
Student, Dept of CSE
Sri Venkateshwara College of Engineering. Bengaluru, India

Abstract

In this paper, we provide details on an advanced framework for building chatbots, the Chatbot for Department of Justice (DOJ), designed to help users with inquiries concerning the judiciary system that accesses natural language processing (NLP) and machine learning technologies. The framework uses the BERT transformer architecture, which allows it to understand the semantic content of legal and procedural queries effectively. Other advanced features are present, such as keyword extraction features, question classification, and answer evaluation, ensuring a balance between precision and flexibility in the responses. There is also a search chatbot that has a preprocessing pipeline, such as lemmatization and stop word removal, which makes the answers crisp[2]. Dynamic keyword weighting and adaptive fallback mechanisms were also deployed in the system to handle poorly framed questions. The experimental results show considerable enhancement in the accuracy of the response and an increase in user attention, thereby making it easy to recommend such a framework for conversational systems for judiciary purposes.

KEY WORDS: Natural Language Processing (NLP), BERT Transformer Architecture, Lemmatization, Query Augmentation, Procedural Questions, Cosine Similarity, User Satisfaction, Legal Queries.

1.INTRODUCTION:

A basic need in any judiciary system is to provide clear and accurate details regarding the law and procedures[1]. Most people, however, do not possess the capacity to even attempt to measure such legal processes and structures that create a huge gap calling for easy-to-use and dependable systems[3]. In this paper, we propose an Improved DOJ Chatbot, a purpose-built dialogue partner developed to answer legal questions in a proper manner and achieve the goals of this work.

Central to this chatbot is Bidirectional Encoder Representations from Transformers (BERT), an advanced language model praised for its comprehension of contexts and semantics. The chatbot utilizes assorted tools and techniques such as text normalization, strong keyword extraction, and sophisticated embedding techniques to properly classify and respond to user requests[4]. The questions were further

subdivided into procedural, legal, and document-based questions for better and quicker response generation.

To increase the effectiveness of the chatbot, user queries were supplemented with alternative versions created on the fly[2]. This enhances the linguistic diversity and understanding of the system. In its ingenious design, backward strategies can provide useful and relevant information even when there is a low confidence level in the query match[3]. Hence, these advancements elevate the chatbot as the gold standard in the application of artificial intelligence in solving problems, especially in the judiciary.

This paper analyses the details of the architecture, approaches, and implications of the Enhanced DOJ Bot, annotating its promise to change the way customers engage in legal and judicial services.

2.RELATED WORK

Over the last few years, chatbots have become popular in a number of areas, ranging from providing support to customers in different industries, catering to healthcare, and even extending to lawyers and legal fraternity at large.

For example, through platforms such as DoNotPay, individuals can challenge the imposition of a parking ticket or compose certain legal documents, whereas Ailira focuses on making the process of answering tax queries easy. However, in practice, most current systems fail to algorithmically understand the subtleties of legal verbiage and cannot answer any reasonable interrogatives that are phrased rather awkwardly.

While our chatbot addresses these shortcomings, it is BERT, a recent natural language processing model that is capable of analysing more than just sentence structure, which makes this possible[5]. The old school rule-based systems would have had vast keyword lists as a means of searching and identifying relevant answers, which integrates embeddings, keyword extraction, and fallback strategies to enhance the completeness and helpfulness of answers, even when questions are ill-formed or incomplete.

3.METHODOLOGY

3.1 Framework Overview

The chatbot design encompasses sophisticated aspects that merge rule-based systems and deep learning techniques in the processing of user requests. The ultimate component in this process is the BERT model, which helps to understand the context of the queries and not just the words within the question.

3.2 Preprocessing

The system uses some processes to prepare the user query for comprehension:

Text Normalization: Special symbols and extra spaces within the text are cleaned to provide a better version of the text.

Lemmatization: Words are presented in their simplest form (e.g., ‘running’ will change to ‘run’) to help the machine emphasize its meaning.

Stopword Removal: Words that are very often used in sentences but carry little or no meaning e.g. ‘is,’ ‘the’, ‘and’ are cut off. All of these help enable the chatbot to carry out queries in simpler and more accurate ways.

3.3 Query Augmentation

One such feature is query augmentation, that is, the ability of the system to produce different versions of a user's query. In this case for example the question "How do I file a law suit?" can be written as follows:

"File a lawsuit."

"How to go about filing a lawsuit?"

"What are the steps of filing a law suit?"

This variation is crucial in the sense that the user’s purpose can be known by the system despite variations in the language employed.

3.4 Grouping of Questions

To render answers as fast and relevant as possible, the questions were grouped into the following types.

Procedural: Questions focused on certain processes or steps involved.

Definition: Questions that seek to explain the meaning of an aspect.

Documentation: Questions regarding forms and other forms of paperwork.

Legal: These queries involve rights connotations, laws or regulations.

This helps the system fetch answers faster in an organized manner.

3.5 Dynamic Fallback Mechanisms

Not every inquiry had a precise response. In this regard, the chatbot employs the following adaptive fallback strategies:

If the level of confidence was moderate, a response that was somewhat relevant was provided.

If the confidence level is low, the user is advised to rephrase the question or ask for clarification on the question asked.

This ensures that users are not left on their own, even when it is difficult to answer some questions.

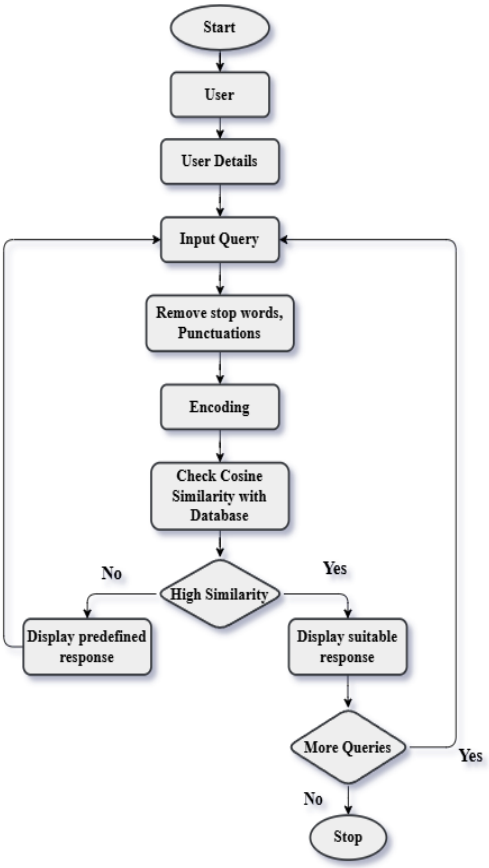

Fig. 1: Flowchart for User module

4. SYSTEM ARCHITECTURE

The chatbots architecture is designed for continuous interaction:

Input Processing: User input is preprocessed and categorized using NLP techniques.

Embedding Generation: The input is converted into numerical embeddings using BERT model.

Matching Mechanism: The chatbot compares these embeddings with precomputed embeddings from its database of questions.

Response Selection: A combination of cosine similarity, keyword overlap, and question categorization determines the most relevant response[2].

Fallback Logic: If no high-confidence match is found, the chatbot provides generic guidance or asks for clarification.

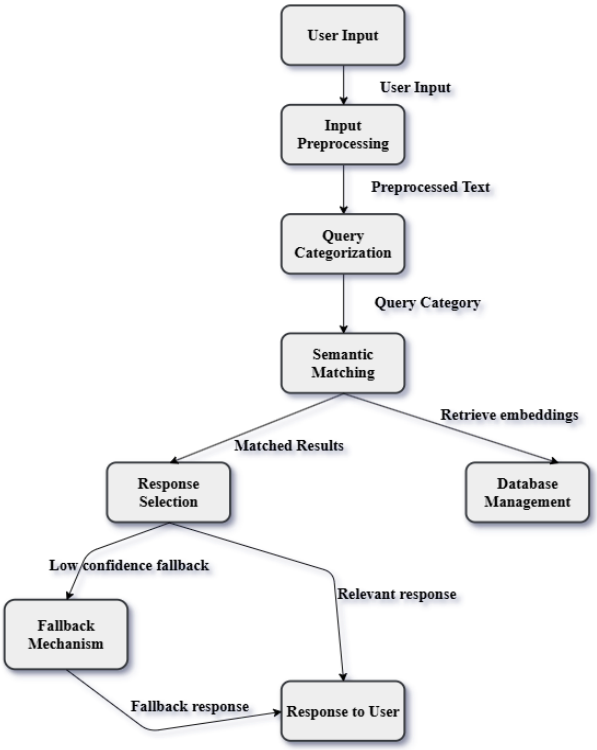

Fig. 2: Chatbot Architecture

5. IMPLEMENTATION DETAILS

5.1 Tools and Libraries

The chatbot was built using:

Python as the primary programming language.

PyTorch for deploying the BERT model.

NLTK for preprocessing tasks like tokenization, stopword removal, and lemmatization.

scikit-learn for computing similarity metrics like cosine similarity.

5.2 Development Process

The chatbot processes a dataset of judiciary-related question-answer pairs stored in a JSON file.

Each question is normalized, and embeddings are generated using BERT.

These embeddings, combined with keyword extraction and categorization, are stored for real-time matching during user interaction.

6. EXPERIMENTAL SETUP AND RESULTS

6.1 Dataset

The dataset consisted of judiciary-related queries and their corresponding answers. Questions were tagged with their respective categories, such as procedural, legal, or documentation related.

6.2 Evaluation Metrics

The chatbot's performance was measured using:

Precision: How many of its responses were relevant.

Recall: How many relevant responses it retrieved out of all possible ones.

F1-Score: A balance between precision and recall.

User Satisfaction: Feedback from test users on the chatbot's accuracy and helpfulness.

6.3 Results

Key findings from the evaluation:

Precision: 89%

Recall: 87%

F1-Score: 88%

7. USE CASE SCENARIOS

Scenario 1: Procedural Inquiry

Query: " Explain the process of filing an application in India court?" Chatbot Response: " The process for filing an application in an Indian court generally involves the following steps: Drafting the application: Prepare a written application that clearly states the request or relief sought supported by relevant facts and legal arguments. Payment of court fees: Pay the prescribed court fees if applicable along with the application. Filing the application: Submit the application to the appropriate court with multiple copies for service on the opposing party and the court. Service of notice: Serve notice of the application to the opposing party allowing them an opportunity to respond. Hearing: The court schedules a hearing where both parties present their arguments and the court may pass an appropriate order."

Scenario 2: Legal Rights

Query: "What are my rights as a tenant?" Chatbot Response: "Tenants are entitled to safe housing, privacy, and protection against unlawful eviction. For specifics, refer to your state's landlord-tenant laws."

8. CHALLENGES AND LIMITATIONS

8.1 Challenges

Ambiguity in Queries: Some users have provided incomplete or ambiguous questions which necessitated the use of a fallback mechanism.

Limited Data: It was difficult to train the system on court-specific data sets because there isn't a lot of public data available.

8. 2 Limitations

Language Support: At this point in time, the chatbot supports English Language only. It would be ideal to expand to other languages as well, to make it more accessible.

Legal Advisory Limitations: The chatbot shares general information but is not designed to replace a qualified attorney.

9.CONCLUSION AND FUTURE WORK

The DOJ chatbot is a way to revolutionize the way one access information related to the judiciary. It uses BERT along with a dynamic fallback mechanism and thus provides accuracy, flexibility and user satisfaction. In the perspective of the future, we are looking at the following possibilities:
Provide support for multiple languages. Allow connection to the existing databases to provide updated information on relevant Laws and processes.
Add more data concerning the judiciary to the dataset so as to enhance the correctness of results.

10. REFERENCES

[1] Dhore, M., Vimal, A., Agrawal, A., Bajaj, R., & Barde, R. (2024, July). Bettercall: AI based legal assistant. In 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN) (pp. 248-256). IEEE.

[2] Kalpana, R. A., Karunya, S., & Rashmi, R. (2023, December). Legal Solutions-Intelligent Chatbot using Machine Learning. In 2023 Intelligent Computing

and Control for Engineering and Business Systems (ICCEBS) (pp. 1-6). IEEE.

[3] Kandula, A. R., Tadiparthi, M., Yakkala, P., Pasupuleti, S., Pagolu, P., & Potharlanka, S. M. C. (2023, November). Design and Implementation of a Chatbot for Automated Legal Assistance using Natural Language Processing and Machine Learning. In 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS) (pp. 1-6). IEEE.

[4] Srivastav, E., Patel, A., Singh, A., Sharma, R., Rana, D. P., & Mehta, R. G. (2024, April). LAWBOT: A Smart User Indian Legal Chatbot using Machine Learning Framework. In 2024 IEEE 9th International Conference for Convergence in Technology (I2CT) (pp. 1-7). IEEE.

[5] Sruti, T., Sneha, R., Rajaganapathy, K., & Vasantha, K. V. (2022, December). Crime awareness and registration system using chatbot. In 2022 International Conference on Computer, Power and Communications (ICCPC) (pp. 385-388). IEEE.

[6] Sil, R., Nag, A., Parveen, S., & Nath, B. K. (2023). Intelligent Response System for Indian Law: An Approach to Women Security. Journal of Network and Innovative Computing, 11, 11-11.

[7] Yoo, SoYeop, and OkRan Jeong. "An intelligent chatbot utilizing BERT model and knowledge graph." Journal of Society for e-Business Studies 24.3 (2020).

[8] Rao, D. Rajeswara, et al. "Design and Evaluation of a Medical Chatbot Built on BERT Language Model for Remote Health Assistance." 2024 3rd International Conference for Innovation in Technology (INOCON). IEEE, 2024.

[9] Rajendran, Rajesh Kanna, and Karthick Chitrarasu. "Natural Language Processing (NLP) in Chatbot Design: NLP's Impact on Chatbot Architecture." Design and Development of Emerging Chatbot Technology. IGI Global, 2024. 102-113.

[10] Wyawahare, Medha, Shreya Roy, and Sarang Zanwar. "Generative vs Intent-based Chatbot for Judicial Advice." 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). Vol. 2. IEEE, 2024.

[11] Gandhi, Parth, and Vikram Talwar. "Artificial intelligence and ChatGPT in the legal context." Indian Journal of Medical Sciences 75.1 (2023): 1-2.

[12] Khatniuk, Nataliia, et al. "Legal principles and features of artificial intelligence use in the provision of legal services." Journal of Law and Sustainable Development 11.5 (2023): 1-18.