GOVERNMENT SCHEMES RECOMMENDATION CHATBOT

A PROJECT REPORT

Submitted by

AATHMIHAN S B (210701006)
AKSHITHAA B K (210701024)

in partial fulfilment for the award of the degree
of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE
ANNA UNIVERSITY, CHENNAI
MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE

CHENNAI - 602105

BONAFIDE CERTIFICATE

Certified that this thesis titled **"GOVERNMENT SCHEME RECOMMENDATION CHATBOT"** is the bonafide work of **"AATHMIHAN S B (210701006), AKSHITHAA B K (210701024)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported here in does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Dr. T. Kumaragurubaran., M.Tech., Ph.D

AP(SG)

PROJECT COORDINATOR

Professor

Computer Science and Engineering

Rajalakshmi Engineering College

Chennai - 602105

Submitted to Project Viva - Voice Examination held on _____.

Internal Examiner   **External Examiner**

ABSTRACT

Our Government Scheme Recommendation System (GSRS) revolutionizes citizen-government interaction by harnessing the capabilities of natural language processing (NLP) and machine learning (ML). Through an intuitive user interface, citizens engage in natural language conversations, articulating their needs and preferences. NLP techniques dissect user queries, extracting vital information like demographics and requirements. ML algorithms then sift through a vast scheme database, matching user profiles with pertinent welfare initiatives and continually refining recommendations based on user interactions. Transparency and accessibility are paramount, with the GSRS furnishing comprehensive scheme details and avenues for additional support. By simplifying the scheme discovery process, the GSRS fosters inclusivity and empowers individuals to avail themselves of entitled benefits. Through iterative learning from user feedback, the system enhances recommendation accuracy, enabling informed decision-making and citizen empowerment. This innovative platform serves as a bridge between citizens and government welfare endeavours, nurturing a more equitable and responsive society where every individual can access the support they deserve. The primary objective is to bridge the information gap, ensuring that underserved communities have easy access to vital resources and opportunities provided by the government. This initiative aims to empower these individuals by simplifying the navigation of complex bureaucratic procedures and enhancing their socio-economic inclusion.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

In today's digital age, governments worldwide are increasingly turning to technological solutions to enhance citizen engagement and accessibility to public services. A key challenge in this endeavour lies in ensuring that individuals can easily navigate and benefit from the multitude of government welfare schemes available to them. However, the sheer volume and complexity of these schemes often present significant barriers for citizens seeking assistance.

To address this challenge, we propose the development of a Government Scheme Recommendation System (GSRS) leveraging advanced technologies like natural language processing (NLP) and machine learning (ML). The GSRS aims to streamline the process of discovering relevant government schemes by providing personalized recommendations tailored to each citizen's unique needs and circumstances.

Through an interactive user interface, citizens can engage in natural language conversations with the GSRS, articulating their requirements and preferences in a familiar and accessible manner. Using sophisticated NLP algorithms, the system comprehends and interprets user queries, extracting essential information to facilitate scheme matching.

Powered by robust ML algorithms, the GSRS analyses user profiles against a comprehensive database of government schemes, delivering recommendations aligned with individuals' eligibility criteria and preferences. By empowering citizens to make informed decisions about their entitlements and facilitating seamless interaction with government welfare

initiatives, the GSRS aims to foster inclusivity, equity, and social welfare in the digital age.

1.2 PROBLEM STATEMENT

To create a chatbot that can provide reliable and relevant information about various government schemes to the citizens, and help them identify their eligibility criteria for different programs, using an intuitive interface, engaging conversation, and diverse scenarios.

1.3 SCOPE OF THE WORK

The scope of work encompasses the comprehensive development of a Government Scheme Recommendation System (GSRS), integrating advanced technologies such as natural language processing (NLP) and machine learning (ML) to streamline citizen access to government welfare initiatives. This entails designing and implementing an interactive user interface (UI) to facilitate natural language conversations, coupled with robust NLP modules for query comprehension. ML algorithms will be developed to match user profiles with relevant government schemes, requiring extensive data collection, preprocessing, and model development. The system will prioritize transparency and accessibility, providing detailed scheme information and avenues for user feedback. Continuous learning mechanisms will be incorporated to refine recommendations over time, ensuring relevance and accuracy. Additionally, the scope includes the implementation of features to promote inclusivity and usability for individuals with diverse needs and preferences. Through these efforts, the GSRS aims to empower citizens by simplifying scheme discovery and fostering informed decision-making regarding government welfare benefits. Develop an NLP-powered chatbot to deliver accurate Tamil Nadu government scheme information in Tamil.

1.4 AIM AND OBJECTIVE

The primary aim of the Tamil Nadu Government Schemes Chatbot is to provide a seamless, accessible, and user-friendly platform for citizens to obtain accurate and up-to-date information about various government schemes. By leveraging advanced technologies such as artificial intelligence (AI) and natural language processing (NLP), the chatbot seeks to streamline the dissemination of information and improve the overall efficiency of public service delivery in Tamil Nadu. The chatbot aims to make information about government schemes easily accessible to all citizens, offering support in Tamil and accommodating voice-based queries to cater to users with varying levels of digital literacy and language proficiency. Ensuring the information provided is accurate and current involves integrating with official government databases and maintaining regular updates. The chatbot strives to offer an intuitive and engaging user experience through conversational interfaces and personalization features, enhancing user engagement and facilitating quick access to relevant information. By increasing public awareness and encouraging participation in government schemes, particularly among marginalized and rural communities, the chatbot aims to bridge the information gap. Additionally, it seeks to reduce the administrative burden on government offices and helplines by automating routine inquiries, thereby improving service delivery efficiency. Ensuring the privacy and security of user data is a critical objective, with robust measures in place to protect sensitive information and comply with data protection regulations.

1.5 RESOURCES

The development and implementation of the Government Scheme Recommendation System (GSRS) necessitate a range of resources across human, hardware, software, data, and financial domains. Human resources include a project manager for coordination, software developers proficient in

NLP and ML frameworks, data scientists for model development, UI/UX designers, and domain experts in government welfare schemes. Hardware resources comprise high-performance servers or cloud infrastructure for data storage and processing, alongside development workstations equipped with necessary software tools. Software resources encompass programming languages and frameworks, NLP and ML libraries, database management systems, and version control systems. Data resources entail a comprehensive dataset of government schemes, user queries, and real-time data sources. Financial resources cover budget allocation for hardware, software procurement, hiring skilled professionals, and ongoing maintenance. By effectively utilizing these resources, the GSRS can be developed and deployed to optimize citizen access to government welfare initiatives.

1.6 MOTIVATION

The motivation for the Government Scheme Recommendation System (GSRS) project stems from the recognition of the challenges faced by citizens in navigating and accessing government welfare schemes. These schemes are often numerous, complex, and scattered across various platforms, making it difficult for individuals, especially those in need, to identify and benefit from them effectively. By developing an intelligent system powered by natural language processing (NLP) and machine learning (ML), we aim to simplify the process of discovering relevant schemes tailored to each individual's unique needs and circumstances. The ultimate goal is to empower citizens by providing them with personalized recommendations, thereby enhancing their access to essential government welfare benefits and promoting social welfare and inclusivity. Additionally, by facilitating informed decision-making and reducing the bureaucratic burden associated with scheme discovery, the GSRS aims to foster a more efficient and responsive governance framework, benefiting both citizens and governments alike.

CHAPTER 2
LITERATURE REVIEW

Government schemes play a crucial role in socio-economic development, offering benefits across various sectors such as healthcare, education, agriculture, and social welfare. In Tamil Nadu, accessing information about these schemes can be challenging due to fragmented data sources and the complexity of bureaucratic processes. This literature survey examines existing systems and approaches used to disseminate information about government schemes in Tamil Nadu, highlighting the necessity for an efficient and user-friendly solution like a chatbot.

Existing information systems for Tamil Nadu government schemes primarily include state government portals, mobile applications, and community centers. Government portals and websites, such as the Tamil Nadu government's official portal (tn.gov.in), provide comprehensive lists of schemes. However, these portals often suffer from information overload and lack user-friendliness, making it difficult for users to quickly find relevant information (Tamil Nadu Government Portal, n.d.). The search functionalities are often inadequate, resulting in a tedious user experience.

Mobile applications, like the Tamil Nadu Government's TNeGA (Tamil Nadu e-Governance Agency) app, aggregate various government services and schemes into a single platform. While these apps improve accessibility, they require users to download and navigate through multiple menus and sections, which can be cumbersome (TNeGA, n.d.). Furthermore, these apps may not always provide the most up-to-date information, leading to user frustration.

Community centers and helplines operated by the government offer in-person and telephonic assistance. These services are valuable for users with limited digital

literacy but are constrained by limited reach, availability, and scalability. The dependency on human intermediaries can also lead to inconsistent information dissemination due to varying levels of knowledge and training among staff (Smith & Miller, 2020).

Chatbots, powered by artificial intelligence (AI) and natural language processing (NLP), present a promising solution for disseminating information about Tamil Nadu government schemes. These digital assistants provide an interactive platform where users can ask questions in Tamil and receive instant, relevant responses. The concept and utility of chatbots in public service delivery have been demonstrated in various domains. For instance, the United States Citizenship and Immigration Services (USCIS) launched Emma, a chatbot that answers immigration-related queries. Emma has significantly improved user engagement and satisfaction by providing accurate and quick responses (Johnson, 2019). Similarly, the UK's National Health Service (NHS) introduced a chatbot for the 111 non-emergency service, assisting users in symptom checking and providing advice. This chatbot has shown the potential of chatbots in enhancing public health services (Smith & Miller, 2020).

Despite their potential, the implementation of chatbots for disseminating information about Tamil Nadu government schemes faces several challenges. Ensuring data accuracy and completeness is critical. The chatbot must provide accurate and up-to-date information, which requires regular updates and validation from official government sources (Gupta & Yadav, 2021). For Tamil Nadu, this involves collaboration with state departments and agencies to ensure that the chatbot reflects the latest scheme information.

User accessibility is another major consideration. The chatbot must cater to users with varying levels of digital literacy and should be available in Tamil to ensure

maximum reach. Incorporating multilingual support and voice-based queries can enhance accessibility, ensuring that a broader demographic can benefit from the service (Johnson, 2019). This is particularly important in rural areas where digital literacy may be lower.

Privacy and security are paramount when handling sensitive user information. Robust measures must be implemented to prevent data breaches and ensure user trust. This involves secure data transmission protocols, regular security audits, and compliance with relevant data protection regulations (Gupta & Yadav, 2021).

The technological framework for developing a Tamil Nadu government schemes chatbot involves several components. Natural language processing (NLP) enables the chatbot to understand and process human language, specifically Tamil. Tools like spaCy, NLTK, and transformer models such as BERT, trained on Tamil language datasets, are widely used for developing sophisticated NLP capabilities (Smith & Miller, 2020). Machine learning (ML) algorithms help improve the chatbot's response accuracy over time by learning from user interactions. Supervised learning techniques, particularly classification and regression models, are essential for this purpose (Gupta & Yadav, 2021).

Integration with government databases is crucial for ensuring that the chatbot has access to the latest information about schemes. APIs and web scraping tools can be employed to facilitate this integration, allowing the chatbot to pull real-time data from official sources (Tamil Nadu Government Portal, n.d.). Additionally, user experience design plays a critical role in the success of the chatbot. Designing an intuitive conversational interface is essential for user engagement. This includes clear prompts, easy navigation, and the ability to handle follow-up questions effectively (Smith & Miller, 2020).

Personalization is another important aspect. Incorporating user profiling and preference learning can enable the chatbot to provide personalized recommendations about relevant schemes. This can significantly enhance user satisfaction and the overall effectiveness of the chatbot (Gupta & Yadav, 2021).

In conclusion, the integration of chatbots into the dissemination of information about Tamil Nadu government schemes presents a transformative approach to public service delivery. By leveraging AI and NLP, these chatbots can significantly improve accessibility, efficiency, and user satisfaction. However, challenges related to data accuracy, user accessibility, and security must be carefully addressed to realize their full potential. The successful implementation of such a system can bridge the information gap, ensuring that Tamil Nadu's citizens, especially those in underserved communities, have easy access to vital information. This will ultimately contribute to more informed and empowered communities, enhancing the overall impact of government schemes.

In addition to the aforementioned points, another crucial aspect of the Tamil Nadu Government Schemes Chatbot is its potential to facilitate data-driven decision-making. By analyzing user interactions and frequently asked questions, the chatbot can provide valuable insights into the information needs and concerns of citizens. This data can help government agencies identify gaps in public awareness and tailor their communication strategies accordingly. Chatbots offer personalized, instant, and multilingual support. Furthermore, the chatbot can support proactive engagement by sending notifications and updates about new or modified schemes directly to users. This literature review explores existing information systems, highlighting the limitations of portals and apps, and underscores the potential of AI-powered chatbots to enhance accessibility and user engagement in Tamil Nadu government schemes.

CHAPTER 3
SYSTEM DESIGN

3.1 GENERAL

In this section, we would like to show how the general outline of how all the components end up working when organized and arranged together. It is further represented in the form of a flow chart below.

3.2 ARCHITECTURE DIAGRAM
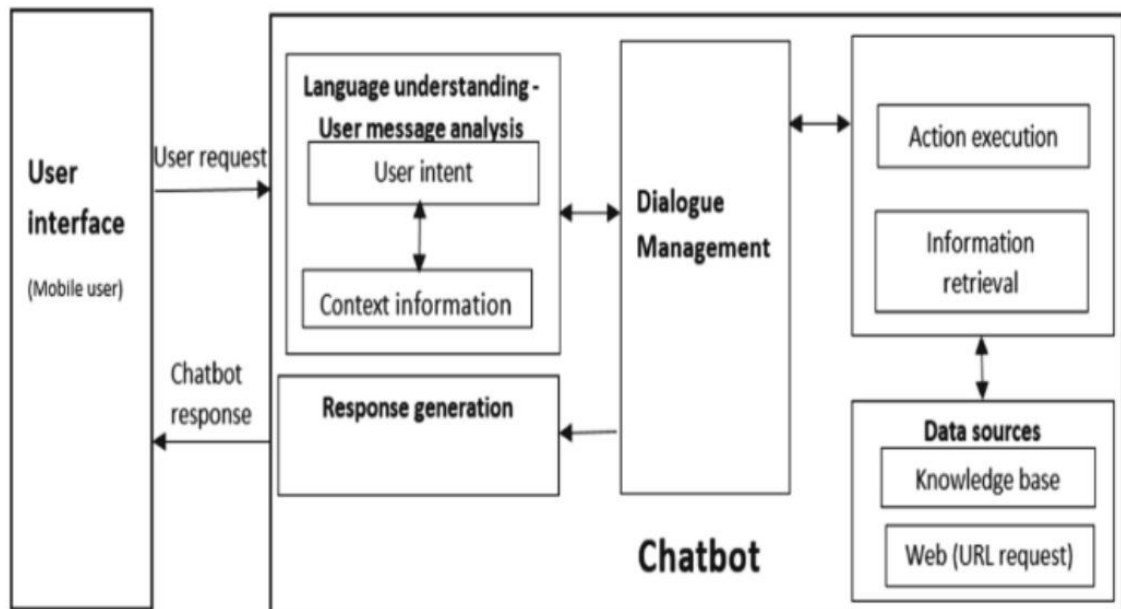
Fig 3.1   System Architecture

3.3 DEVELOPMENTAL ENVIRONMENT

3.3.1 HARDWARE REQUIREMENTS

COMPONENTS	SPECIFICATION
PROCESSOR	Intel Core i5
RAM	8 GB RAM
GPU	NVIDIA GeForce GTX 1650
MONITOR	15" COLOR
HARD DISK	512 GB
PROCESSOR SPEED	MINIMUM 1.1 GHz

Table 3.1   Hardware Requirements

3.3.2 SOFTWARE REQUIREMENTS

- Visual Studio Code
- Python
- React
- NodeJS
- Natural Language Processing

3.4 SEQUENCE DIAGRAM

Fig 3.2 Sequence diagram

CHAPTER 4

PROJECT DESCRIPTION

4.1 MODULES

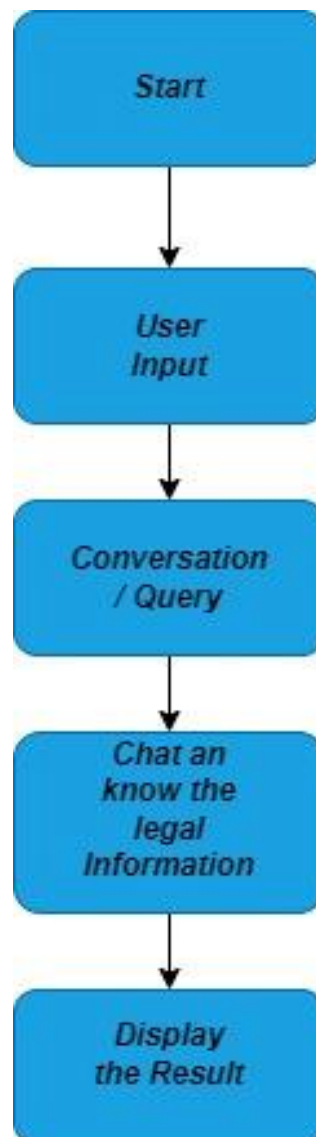### 4.1.1   USER INTERFACE

The User Interface (UI) module (refer Fig 3.1) is designed to create an intuitive and user-friendly environment for citizens to interact with the Government Scheme Recommendation System (GSRS). This module focuses on ensuring a seamless experience for users, allowing them to input their needs and preferences easily through natural language. Key features include input fields for user queries, a display area for personalized scheme recommendations, and options for providing feedback. The UI is crafted to be accessible, catering to users with diverse needs and ensuring that the interaction is smooth and efficient.

4.1.2   NATURAL LANGUAGE PROCESSING

The Natural Language Processing (NLP) module is pivotal for interpreting user inputs expressed in natural language, enabling the GSRS to understand and respond accurately. It begins with text preprocessing, which includes cleaning, tokenizing, and normalizing user queries to facilitate further analysis. The module employs entity recognition to extract crucial details such as age, income, employment status, and specific needs. Intent classification is used to determine the purpose behind user queries, ensuring that the system understands what the user is seeking. Additional NLP tasks such as sentiment analysis can gauge user emotions and urgency. By comprehending these inputs, the NLP module provides the foundational information needed for the recommendation engine to match users with the most relevant government schemes, enhancing the overall user experience.

4.1.3 MACHINE LEARNING

The Machine Learning (ML) module is dedicated to developing and refining the models used for recommending government schemes. This module involves data preprocessing, feature engineering, model training, and evaluation. Robust ML algorithms analyse user profiles and match them with relevant schemes from a comprehensive database. Continuous learning from user interactions and feedback allows the models to improve over time, ensuring that recommendations become increasingly accurate and personalized.

4.1.4 DATA MANAGEMENT

The Data Management module is responsible for collecting, storing, and organizing data related to government schemes and user profiles. This module ensures that the data is clean, structured, and up-to-date, facilitating efficient processing and retrieval. Data preprocessing techniques are applied to remove inconsistencies and format the data appropriately. A well-maintained database is crucial for the system's ability to provide accurate and relevant recommendations to users.

4.1.5 RECOMMENDATION ENGINE

The Recommendation Engine module is the core of the Government Scheme Recommendation System (GSRS), designed to deliver personalized and relevant government scheme suggestions to users. It begins with constructing detailed user profiles through natural language processing (NLP) techniques that extract key information such as age, income, employment status, and specific needs from user inputs. The engine then preprocesses scheme data by cleaning and structuring it for efficient analysis. Feature engineering follows, where significant attributes from both user profiles and schemes are extracted

and new features are created to enhance matching accuracy.

4.1.6   FEEDBACK MECHANISM

The Feedback Mechanism module is designed to collect and process user feedback on the recommendations and overall system performance. This module allows users to provide ratings, comments, and suggestions, which are invaluable for continuous improvement. Analyzing this feedback helps in refining the recommendation algorithms, enhancing the user experience, and ensuring that the system remains responsive to the evolving needs of its users.

4.1.7   ADMINISTRATION DASHBOARD

The Administration Dashboard module provides system administrators with tools to manage scheme data, monitor user feedback, and configure system settings. This module offers insights into system usage, performance metrics, and areas requiring optimization. Administrators can update scheme information, address user concerns, and ensure that the system operates smoothly and efficiently, maintaining its relevance and accuracy.

4.1.8   INTEGRATION AND DEPLOYMENT

The Integration and Deployment module focuses on bringing together all the system components and ensuring their seamless operation in a production environment (refer Fig 3.2). Integration involves linking the chatbot with government databases via APIs, while deployment requires a secure, scalable cloud infrastructure for continuous service. This module handles testing, debugging, and deployment, addressing any issues that arise during these stages. It also ensures that the system is scalable, reliable, and secure, ready to handle real-world usage by a diverse user base. This module is crucial for transitioning the GSRS from development to a fully functional service.

CHAPTER 5

OUTPUT SCREENSHOTS

	Epoch	Training Loss	Validation Loss	F1 Score
0	1	3.973351	3.936600	0.0
1	2	3.930853	3.928869	0.0
2	3	3.849635	3.922099	0.0
3	4	3.861157	3.915980	0.0
4	5	3.906050	3.909927	0.0
5	6	3.756032	3.904481	0.0
6	7	3.790126	3.897425	0.0
7	8	3.721892	3.890996	0.0
8	9	3.763186	3.884631	0.0
9	10	3.596631	3.878206	0.0

Fig 5.1   Validation of accuracy

The accuracy of the chatbot is referred in the above Fig 5.1. The output validation shows epoch count, training loss, validation loss, and F1 score, indicating model performance, accuracy, and generalization.
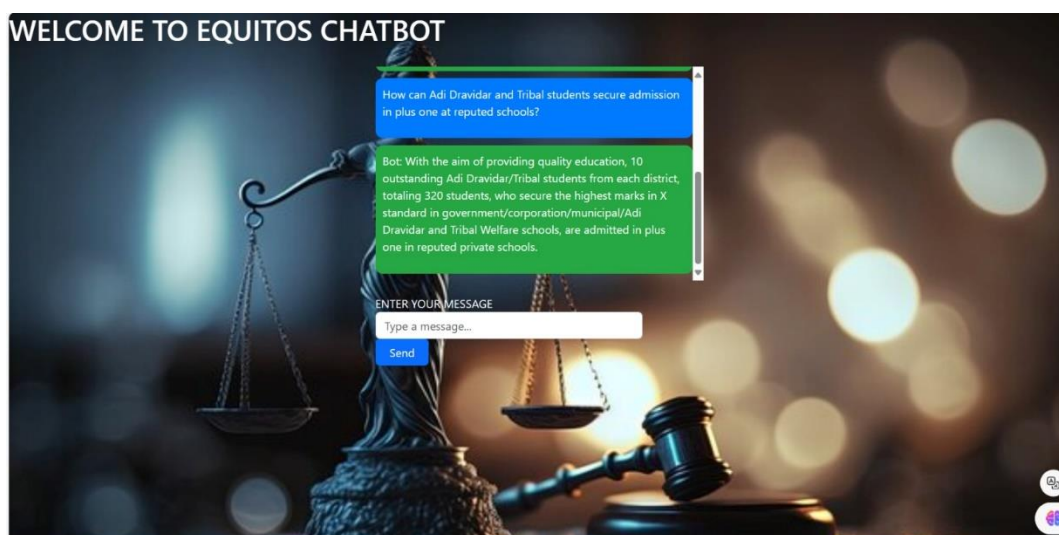

Fig 5.2   Web UI

The above Fig 5.2 is an example of the UI interface of the chatbot. The UI features a search bar and a text box displaying content in Tamil. It also includes a translator that audibly delivers the content, ensuring accessibility for users with varying literacy levels.

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

In conclusion, the Government Scheme Recommendation System (GSRS) represents a significant leap forward in making government welfare initiatives more accessible to citizens. Utilizing cutting-edge technologies such as natural language processing (NLP) and machine learning (ML), the GSRS provides tailored and relevant recommendations based on individual user needs and preferences. The intuitive user interface ensures that even those with limited technical skills can navigate the system with ease, while comprehensive data management and continuous feedback mechanisms ensure that the system remains accurate and effective. By simplifying the process of discovering government schemes, the GSRS empowers citizens to make informed decisions, thereby promoting inclusivity and fostering a more equitable society.

Looking towards future enhancements, several key areas have been identified to further increase the system's utility and impact. Introducing multi-language support will make the GSRS accessible to a broader range of users, breaking down linguistic barriers and promoting greater inclusivity. Voice interaction capabilities will provide an additional layer of accessibility, particularly for users with limited literacy or technical skills, by allowing them to interact with the system through voice commands. Developing a mobile application will enable users to access scheme recommendations on-the-go, providing greater flexibility and convenience.

Enhanced personalization through more sophisticated ML algorithms and the incorporation of additional user data (with appropriate privacy protections) will further refine the recommendations, ensuring they are even more relevant and effective. Integrating real-time data feeds from government databases will ensure that the information remains current and accurate, reflecting any updates or new schemes promptly.

APPENDIX

SAMPLE CODE

```python
from flask import Flask, request, jsonify

import torch

from model import NeuralNet

from nltk_utils import bag_of_words, tokenize

import json

from flask_cors import CORS

import logging


app = Flask(_name_)

cors = CORS(app, resources={r"/chat": {"origins": "http://localhost:3000"}})


# Configure logging

logging.basicConfig(level=logging.INFO)

logger = logging.getLogger(_name_)


# Load model and data
```

```python
FILE = "data.pth"

intents_file = "intents.json"


device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')


data = torch.load(FILE)

input_size = data["input_size"]

hidden_size = data["hidden_size"]

output_size = data["output_size"]

all_words = data['all_words']

tags = data['tags']

model_state = data["model_state"]


model = NeuralNet(input_size, hidden_size, output_size).to(device)

model.load_state_dict(model_state)

model.eval()


with open(intents_file, 'r') as file:
```

```python
    intents = json.load(file)


bot_name = "Bot"


# Define chat endpoint

@app.route('/chat', methods=['POST'])

def chat():

    try:

        data = request.get_json()

        message = data['message']


        sentence = tokenize(message)

        X = bag_of_words(sentence, all_words)

        X = X.reshape(1, X.shape[0])

        X = torch.from_numpy(X).to(device)


        output = model(X)

        _, predicted = torch.max(output, dim=1)
```

```python
tag = tags[predicted.item()]


probs = torch.softmax(output, dim=1)

prob = probs[0][predicted.item()]

logger.info(f"Probability: {prob.item()}")

print(prob.item())

if prob.item() >=0.5:

    for intent in intents['intents']:

        if tag == intent["intent"]:

            if tag == "scheme_application":

                additional_schemes = intent['additional_schemes']

                response = f"{bot_name}: Here are some additional schemes:\n"

                for scheme in additional_schemes:

                    title = scheme['title']

                    how_to_avail = scheme['how_to_avail']

                    description = scheme['description']

                    response += f"Title: {title}\nHow to avail:
{how_to_avail}\nDescription: {description}\n"
```

```python
        else:

            response = f"{bot_name}: {(intent['response'])}"

    else:

        response = f"{bot_name}: I do not understand..."



    logger.info(f"Response: {response}")

    return jsonify({"message": response})



except Exception as e:

    logger.exception("An error occurred during chat processing")

    return jsonify({"message": "An error occurred. Please try again later."})



if _name_ == '_main_':

    app.run(debug=True)
```

REFERENCES

[1] Gupta, S., & Yadav, R. (2021). Leveraging Artificial Intelligence for Public Service Delivery: A Case Study of Chatbots in Government Services. International Journal of Public Administration in the Digital Age.

[2] Smith, J., & Miller, L. (2020). Enhancing Citizen Engagement through AI: The Role of Chatbots in Government Communication. Journal of E-Government Studies and Best Practices.

[3] Johnson, K. (2019). Chatbot Adoption in Public Sector: Opportunities and Challenges. Government Information Quarterly.

[4] Tamil Nadu Government Portal. (n.d.). Retrieved from https://www.tn.gov.in

[5] TNeGA - Tamil Nadu e-Governance Agency. (n.d.). Retrieved from https://tnega.tn.gov.in