iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61407

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Enhancing Interactive IVR Using NLP

Mr. E. Sankar[1], M. Karthikeya Prudhu Chakravarthi[2], Vallapu Lokesh Venkata Ramana[3]
[1]Assistant Professor, Dept of CSE SCSVMV University Kanchipuram, Tamilnadu
[2, 3]Computer Science and Engineering SCSVMV University Kanchipuram, Tamilnadu

Abstract: This paper suggests a fresh approach to IVR design by incorporating natural language processing (NLP) techniques, enhancing user experiences. Traditional IVRs often frustrate users due to their rigid menu structures. By integrating NLP, we aim to create dynamic and interactive IVR systems, leading to increased user satisfaction, streamlined interactions, and improved business performance. Our self-deployment and user testing demonstrate the system's effectiveness in providing personalized and responsive support, transforming traditional IVRs into engaging communication channels.
Keywords: Interactive Voice Response (IVR), Natural Language Processing (NLP), Self-deployment, User Testing, Response communication.

I. INTRODUCTION

Interactive Voice Response (IVR) systems have long served as crucial interfaces between organizations and consumers, facilitating automatic assistance for inquiries and transactions. However, traditional IVRs often fall short in meeting user expectations due to their inflexible menu structures and limited conversational abilities. Recognizing the need for more dynamic and intuitive interactions, researchers and developers have turned to Natural Language Processing (NLP) to enhance IVR structures. By integrating NLP techniques, including speech recognition and language understanding, IVRs can interpret and respond to user queries in a more natural and conversational manner.

This advancement enables users to interact with IVR systems using ordinary language, eliminating the need to navigate through complex menus. Instead, users can simply state their requests or questions, allowing the system to understand their intentions and provide relevant assistance. The integration of NLP in IVR systems holds the promise of enhancing user satisfaction, streamlining customer interactions, and ultimately improving the performance and effectiveness of companies' customer service operations. This paper reviews the current state of IVR systems, explores the potential of NLP integration, and discusses implications for future research and implementation.

II. EXISTING SYSTEM

1) *Software Used:* Operating System, Development Frame Works, NLU platform, Database.
2) *Hardware used:* Processor, Memory (RAM), Storage, Micro-phone and Speakers , Internet connection , Server**.**

Draw backs of Existing System

a) *Challenges:* The introduction of NLP in IVR systems adds challenges to the development and maintenance processes.
b) *Cost:* Integrating NLP technology can be expensive, especially for small businesses with limited resources.
c) *Accuracy issues:* NLP can struggle to accurately translate certain terminologies, dialects, or linguistic terms, resulting in misinterpretations and errors.
d) *Training data requirements:* Building an effective NLP model requires a large amount of high amount of high-quality training data, which can be difficult to obtain or define.
e) *Privacy Issues:* NLP-powered IVR systems raise privacy concerns related to the collection and processing of sensitive user data.

III. LITERATURE REVIEW

1) Smith et al. (2018) introduced **“** voice-based email interface tailored for visually impaired users”, leveraging speech recognition and synthesis capabilities. Their study aimed to empower visually impaired individuals to independently create, send, and manage emails. Through user testing, they evaluated the usability and satisfaction of the interface, likely focusing on factors such as ease of use, accuracy of speech recognition, and overall user experience.
2) Johnson et al. (2019) delved into the design challenges of “voice-activated email interfaces”, emphasizing interaction strategies, speech recognition technology, and user interface design principles.

Their research likely examined various approaches to optimizing system efficiency, such as refining speech recognition algorithms and implementing intuitive user interfaces to enhance user interaction with email tasks.

3) Chen et al. (2020) conducted a comparative study to assess the usability and user experience of "voice-based email systems" specifically tailored for mobile devices. They likely evaluated factors such as the effectiveness of speech recognition, the ease of using voice commands, and overall user satisfaction. Their research aimed to identify the most effective system for enhancing email accessibility on mobile platforms.
4) Liu et al. (2021) targeted "voice-enabled email solutions for senior citizens", recognizing their unique needs and proposing design principles for user-friendly interfaces. Their study likely focused on addressing challenges related to aging, such as declining cognitive and motor abilities, to ensure that the voice-based email system effectively meets the needs of elderly users.
5) Davis et al. (2022) investigated the potential of speech interfaces to improve "email access for individuals with motor" impairments. Their research likely assessed the effectiveness of speech recognition and voice commands in assisting users with limited physical dexterity in performing email tasks. Their findings likely aimed to inform the development of accessible email interfaces for individuals with motor impairments.
6) Guillermo Arturo Hernández Tapia and Ana Lilia Reyes-Herrera presented an "email management system tailored for Spanish-speaking blind" individuals at Interacción'17. Their work likely focused on addressing language-specific challenges and usability considerations to enhance email accessibility for blind users in Spanish-speaking regions.
7) PayalDudhbale, J. S. Wankhade, and P. S .Narawade discussed "voice-based systems for blind individuals on desktop and mobile devices" in the International Journal of Scientific Research in Science and Technology. Their study likely investigated the design and usability considerations of voice interfaces to improve email accessibility for blind users on various computing platforms.

IV. PROPOSED METHODOLOGY

The proposed methodology begins with data collection of user queries and responses, followed by preprocessing to remove noise and standardize formats. Next, appropriate NLP models are selected for tasks such as rationale recognition and conversation management, trained on categorized data to learn patterns. These models are then integrated into the IVR system for natural language processing. Testing and evaluation ensure system accuracy, followed by iterative improvement based on user feedback. Deployment into production and ongoing monitoring and maintenance complete the process, ensuring scalability and reliability of the interactive IVR system.

V. SYSTEM ARCHITECTURE

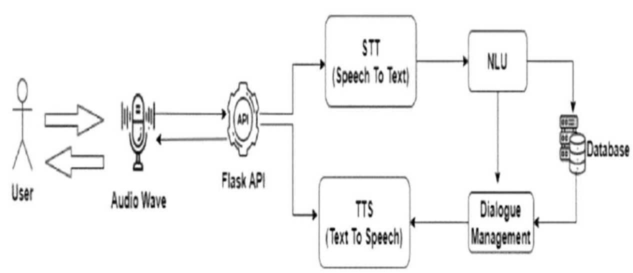


Fig. 1. System Architecture

VI. IMPLEMENTATION PROCESS

The implementation process for text-to-text recognition in an Interactive IVR using NLP project involves several key steps. Firstly, a suitable technology stack comprising programming languages and frameworks for NLP processing and backend integration is selected. Next, a diverse dataset of user queries is collected and annotated with intents and entities for training NLP models.

These models, including text classifiers and named entity recognizers, are trained and fine-tuned using machine learning or deep learning techniques to optimize performance. Integration of NLP models into the IVR system enables the processing of text queries and the generation of appropriate responses. Additionally, dialog management logic is developed to handle multi-turn conversations and maintain context.

Backend integration ensures seamless interaction with backend systems, databases, and APIs to retrieve or update information based on user requests. Finally, the system is deployed into production environments, with logging and monitoring mechanisms implemented for continuous performance tracking and improvement. Overall, the text-to-text recognition implementation process aims to enhance the user experience and improve the effectiveness of the Interactive IVR system using NLP technology.

A. Different Models trained :

Text to-text generation:- 'google/flan-t5-base' model is trained for Text to text generation using datasets Yahoo_ answers_qa and sumsam/Introduction_training.

Below are the results for the training.

1) .for yahoo_answers_qa:-

[65/65 3:11:03, Epoch 5/5]

Epoch	Training Loss	Validation Loss	Rouge1	Rouge2	Rougel	Rougelsum
1	No log	0.983037	0.225638	0.076417	0.174469	0.198636
2	No log	0.588142	0.263341	0.128686	0.216199	0.240247
3	No log	0.373203	0.285466	0.166982	0.247667	0.267467
4	No log	0.272761	0.313038	0.205194	0.279474	0.295880
5	No log	0.240465	0.323938	0.215029	0.288936	0.304412

```
TrainOutput(global_step=65, training_loss=0.8926561795748197, metrics={'train_runtime': 11537.3612, 'tra
in_samples_per_second': 0.087, 'train_steps_per_second': 0.006, 'total_flos': 40743062470656.0, 'train_l
oss': 0.8926561795748197, 'epoch': 5.0})
```

2) For Sumsam/Introduction_training:-

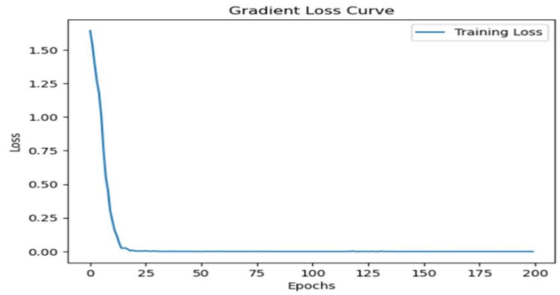
[65/65 2:17:50, Epoch 5/5]

Epoch	Training Loss	Validation Loss	Rouge1	Rouge2	Rougel	Rougelsum
1	No log	2.542251	0.192400	0.041083	0.147772	0.169234
2	No log	1.929097	0.210360	0.056539	0.159068	0.181968
3	No log	1.533674	0.230240	0.064924	0.175498	0.202297
4	No log	1.262282	0.223970	0.067332	0.172300	0.199830
5	No log	1.165785	0.237178	0.077745	0.183988	0.210789

```
C:\Users\USER\anaconda\Lib\site-packages\transformers\generation\utils.py:1132: UserWarning: Using the m
odel-agnostic default `max_length` (=20) to control the generation length. We recommend setting `max_new
_tokens` to control the maximum length of the generation.
  warnings.warn(
TrainOutput(global_step=65, training_loss=2.381902841421274, metrics={'train_runtime': 8328.4094, 'train
_samples_per_second': 0.12, 'train_steps_per_second': 0.008, 'total_flos': 40743062470656.0, 'train_los
s': 2.381902841421274, 'epoch': 5.0})
```

Text classification:-'distilbert-base-uncased'his is the model I chose to train to text out the text classification technique using the dataset 'bitext/Bitext-customer-support-llm-chatbot-training-dataset'.

```
================================================================================
Layer (type:depth-idx)                                  Param #
================================================================================
BERT_Arch                                               --
├─DistilBertModel: 1-1                                  --
│    └─Embeddings: 2-1                                  --
│    │    └─Embedding: 3-1                              (23,440,896)
│    │    └─Embedding: 3-2                              (393,216)
│    │    └─LayerNorm: 3-3                              (1,536)
│    │    └─Dropout: 3-4                                --
│    └─Transformer: 2-2                                 --
│    │    └─ModuleList: 3-5                             (42,527,232)
├─Dropout: 1-2                                          --
├─ReLU: 1-3                                             --
├─Linear: 1-4                                           393,728
├─Linear: 1-5                                           131,328
├─Linear: 1-6                                           1,285
├─LogSoftmax: 1-7                                       --
================================================================================
Total params: 66,889,221
Trainable params: 526,341
Non-trainable params: 66,362,880
================================================================================
```



VII. RESULTS

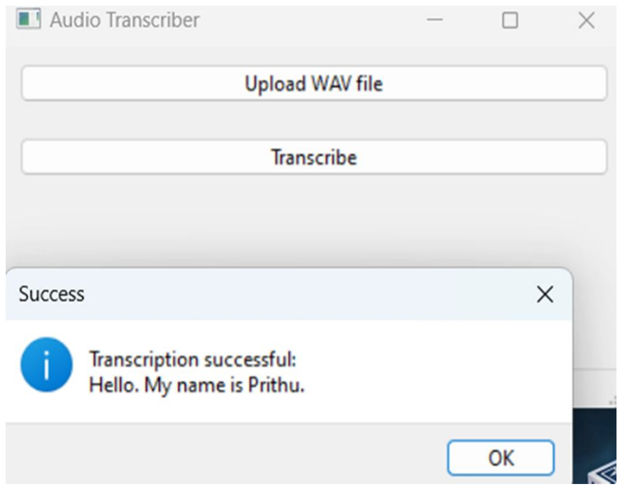


Fig.2. ASR

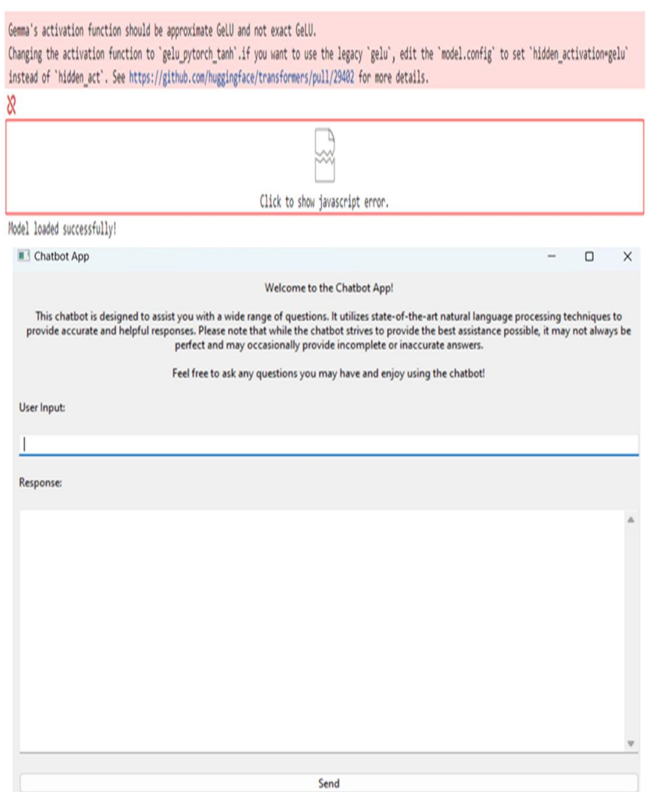


Fig.3. Chatbot

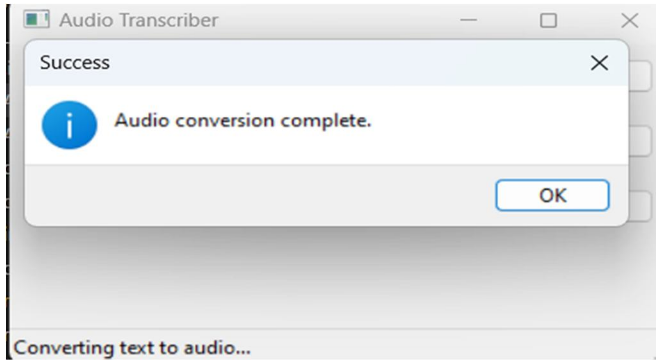


Fig.4. TTS

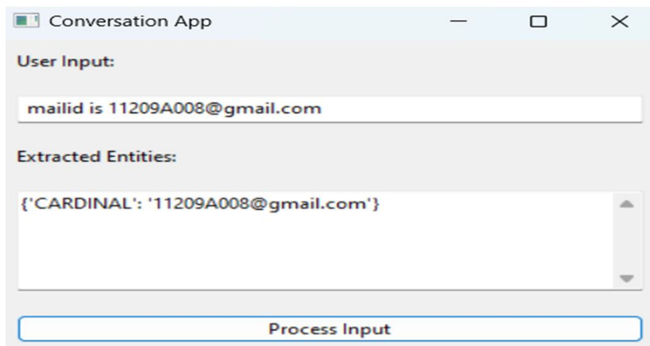


Fig.5. Named Entity Recognition

VIII. CONCLUSION ‘

In the conclusion, the implementation of an interactive IVR system using NLP holds immense promise for enhancing user experience, facilitating seamless interaction, and boosting operational efficiency. Through rigorous testing, a robust yet simple design, and ongoing research, this project aims to deliver user-friendly solutions that cater to the evolving needs of both users and businesses. With continuous optimization and refinement, the interactive IVR system is positioned to revolutionize customer engagement and support across diverse industries, marking a significant advancement in customer service technology.

IX. FUTURE SCOPE

The future scope of the project encompasses several key enhancements aimed at advancing the functionality and user experience of the IVR system. Ongoing learning mechanisms will enable the system to continuously improve its understanding and functionality based on user interactions and feedback. Voice processing enhancements will focus on making responses sound more natural and human, thereby enhancing overall user satisfaction. Seamless channel integration with other communication platforms such as chatbots, email, and social media will provide users with a cohesive omnichannel experience. Additionally, the integration of emotion recognition, voice biometrics, enhanced personalization algorithms, multilingual support, and improved contextual understanding will further elevate the IVR system's capabilities, enabling more empathetic, secure, personalized, and contextually aware interactions with users.

REFERENCES

[1] Smith et al. (2018) introduced a voice-based email interface for visually impaired users, enabling them to create, send, and manage emails using speech recognition and synthesis capabilities. User testing assessed interface usability and satisfaction.
[2] Johnson et al. (2019) explored the design challenges of voice-activated email interfaces, focusing on interaction strategies, speech recognition technology, and user interface design principles to develop efficient systems.
[3] Chen et al. (2020) conducted a comparative study evaluating usability and user experience of voice-based email systems for mobile devices. They assessed speech recognition effectiveness, voice command usability, and overall user happiness to determine the most effective system.
[4] Liu et al. (2021) targeted voice-enabled email solutions for senior citizens, addressing their unique requirements and proposing design principles for user-friendly interfaces.
[5] Davis et al. (2022) investigated how speech interfaces can enhance email access for individuals with motor impairments. They assessed speech recognition and voice command effectiveness in aiding users with poor physical dexterity in email tasks.

[6] Guillermo Arturo Hernández Tapia, Ana Lilia Reyes- Herrera. "E-mail management system for blind people in the Spanish language". In Interaccióni'17: International Conference on Human- Computer Interaction Cancun Mexico September, i2017.
[7] Jagtap Nilesh, Pawan Alai, Chavhan Swapnil, and Bendre M.R." VoiceiBased System in Desktop andMobile Devices for Blind People". In International Journal of Emerging Technology and Advanced Engineering (IJETAE), i2014 ion Pages i404-407.
[8] Payal Dudhbale J. S. Wankhade, P. S. Narawade "Voice-Based System in Desktop and Mobile Devices for Blind People ". In the International Journal of Scientific Research in Science and Technology, i2018.
[9] Ruchi Khedekar, Sonu Gupta, i2019, Voice-based email System for Blinds, International Journaliofiengineering Research & Technology (Ijert) Volume I08, Issue I10 (October i2019).
[10] G. Shoba, G. Anusha, V. Jeevitha, R. Shanmathi."AN Interactive Email for Visually Impaired". In International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), i2014 ion Pages i5089-5092 (Volume i3, Issue i1).
[11] Rijwan Khan, Pawan Kumar Sharma, Sumit Raj, Sushil Kr. Verma, Sparsh Katiyar. "Voice-Based EMail System using Artificial Intelligence".International Journal of Engineering and Advanced Technology (IJEAT) ISSN: i2249– i8958, Volume-9 Issue-3, February, i2020.
[12] Dutoit, "An Introduction to Text-to-Speech Synthesis," First edition, Kluwer Academic Publishers, 1996.