Artificial Intelligence based Virtual Personal Assistant

Dahiwal Aishwarya[1], Gadge Ankita[2], Deokar Manasi[3]

[1-3]Department of Computer Engineering, JCOE, Pune, Maharashtra

---***---

Abstract-*As we know that the Artificial Intelligence is widely occupying the area of technologies which are used in human life, which are affiliated by the affiliated by the appearance and dissemination of internet of things (IOT). This applications are also solves the real world problem. we have chosen this topic for our implementation of final year project which is personal virtual voice assistant. One of the trending fact is that we can use this technology of artificial intelligence for recognizing the Natural Language of human. New insight of this topic can lead to new means of natural human-machine interaction by giving the accurate results. The main aim of this project is that we are going to described which are the principles we have used for developing that virtual personal assistant, how it will work in future enhancement and what are the outcomes we have achieved.*

Key Words: NLP, NLU, TTS, NER, ASR, Internet of things, Virtual Assistant, smart things, Artificial Intelligence.

1.INTRODUCTION

Today the development of artificial intelligence based devices which are able to interact with human by understanding their need at the time of user want (through voice, communication, gestures, facial expressions, etc.) are occurring popularity. A virtual assistant is a digital assistant that uses voice recognition, language processing algorithm, and voice synthesis to listen to specific voice commands and return relevant information or perform specific task as requested by the user. Voice assistant are integrated into many of the devices such as cell phones, computers, smart speakers .Because of this wide array of integrations , there are several virtual voice assistants who offer a very specific feature set, while some choose to be open ended with almost any situation at hand.

2. MOTIVATION

For most of us, the ultimate luxury would be as assistant who always listens for your call, anticipates our every need, and takes action when needed. Also technology is constantly advancing and virtual assistant market will progress a lot along with it.so we are going to build a virtual voice assistant that is accessible in no only in one language but also in other languages as well as it will reminds you important things on accurate situation or location.

3. LITERATURE SURVEY

Speech recognition features a long history with many waves of major innovations. Speech recognition for dictation, search, and voice commands has become a typical feature on smartphones and wearable devices. Design of compact massive vocabulary speech recognition system which will run with efficiency on mobile devices, accurately and low latency. [1] this is often achieved by employing a CTC-based LSTM acoustic model that predicts context-independent phones and is compressed to a tenth of its original size employing a combination of SVD based mostly compression and division. Quantity deep neural networks (DNNs) and on the fly language model rescoring to attain time period performance on fashionable good phones.[2] The ASR and search parts perform speech recognition and search tasks. Additionally to ASR and search, we have a tendency to additionally integrate a question parsing module between ASR and look for variety of reasons. [3] Set of techniques for up the performance of machine-driven voice search services meant for mobile users accessing these services over a spread of moveable devices. Voice search is enforced as a 2 stage search procedure wherever string candidates generated by associate degree automatic speech recognition (ASR) system square measure rescored so as to spot the most effective matching entry from a probably terribly massive application specific info. Study provides a decent example of however further domain specific information sources will be used with a website freelance ASR system to facilitate voice access to on-line search indices. Before AI we tend to were those United Nations agency were upgrading technology to try and do a task however currently the machine is itself able to counter new task and solve it while not ought to involve the humans to evolve it. This has been useful in day to day life-style. From mobile phones to private desktops to mechanical industries these assistants square measure in substantially demand for automating tasks and increasing potency [4] In 2018, fourth world conference on good trends in system, security and property has enforced the thought of creating human interaction with technology easier and a lot of economical. However they found the matter of delayed activation at the incorrect time.

In 2020 , activity user expertise quality of voice assistant has enforced the tactic that relies on the UEQ+ Framework and that permits the measuring of uxor by combining totally different uxor aspects. However they found the matter of inaccurate recognition of speech. In June 2020, domain specific intelligent personal assistant with bilingual voice command process was developed by TENCON. They need created refined technology that may automatise a lot of processes and tasks. However they found the matter for the generalizer, 2 world might have same spellings however totally different meanings. It are often a problem once scaling up.

4. METHODOLOGY

4.1. Front end user

User are having direct access to the interface and communication users by providing Input and Output with graphics styles and icon-based menu. It receives user prompt input voice and reciprocally delivers users with a voice recognition system to sight voice inputs, and frequently generates feedback of voice to users, when completion of commands by many different functions of the system.

4.2 .End user

Basically finish users refers to device users. they'll be mistreatment this devices for communication and feedback of voice with the employment of application, and furthermore finish users area unit people who are mistreatment this application with there personal devices like mobile and portable computer users.
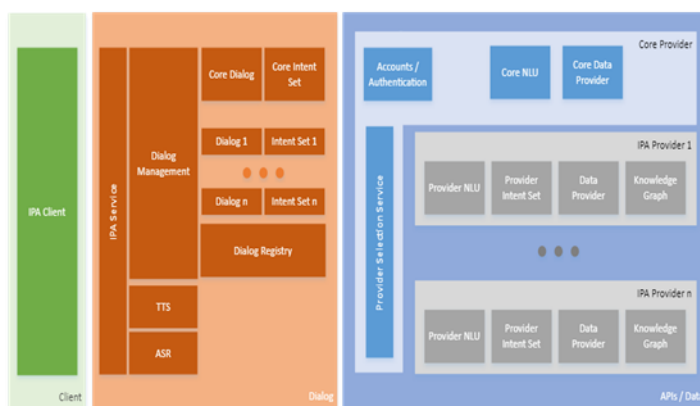
5. SYSTEM ARCHITECURE

Fig-1 . System Architecture

In order to address such use cases as delineated on top of, Associate in Nursing IPA might have to form use of many services describing the capabilities of the IPA. These services could also be chosen from a uniform marketplace. For the reminder of this document, we tend to contemplate Associate in Nursing IPA that's long via such a t place. this sort of IPA options the discipline buildings blocks shown within the on top of figure.

This design contains three layers that area unit careful within the following sections:

Shopper Layer

Dialog Layer

APIs/Data Layer

Actual implementations might want to tell apart quite these layers.

5.1 shopper layer

5.1.1 IPA shopper

Clients alter the user to access the IPA via voice with the subsequent characteristics.

• Usually, IPA shoppers build use of a mike to capture the spoken input and a loud speaker to supply responses.

• As Associate in Nursing extension IPA shoppers can also capture input via text and output text.

• As Associate in Nursing extension IPA shoppers can also capture input from specific modality recognizer.

5.2 Dialog Layer

5.2.1 IPA service

General IPA Service API that mediates between the user and overall IPA system. The service layer could also be omitted just in case the IPA shopper communicates directly with Dialog management. However, this is often not suggested because it might contradict the principle of separation-of-concerns

5.2.2 Dialog Management

Component that receives linguistics info determined from user input, updates its internal state, decides upon consequent steps to continue a dialog and provides output in the main as synthesized or recorded utterances. it's following characteristics

• Dialog Management receives recorded voice input from the IPA services and forwards it to the ASR

• Dialog Management makes use of the TTS to come up with audio knowledge to be rendered on the IPA consumer.

5.2.3 ASR

The automatic Speech Recognizer (ASR) receives audio streams of recorded utterances and generates a recognition hypothesis as text strings. Conceptually, ASR may be a modality recognizer for speech.

5.2.4 TTS

The Text-to-Speech (TTS) element receives text strings, that it converts into audio knowledge. Conceptually, TTS may be a modality specific renderer for speech.

5.2.5 Core Dialog

The Core Dialog is in a position to handle basic practicality via Core Intent Sets to alter interaction with user the least bit. This includes among others

• Greetings

• Goodbye

• Exception handling just in case a requested service isn't accessible

• Exception handling just in case a requested intent cannot be matched to a renowned Dialog

• Help

5.2.6 Core Intent Sets

A Core intent set typically identifies task to be dead and outline the capabilities of the Core Dialog. Conceptually, the Core Intent Sets square measure intent sets that square measure continuously out there.

5.2.6.1 Intent Sets

• An Intent Set defines one or additional intents with AN elective range (including none) of entities to satisfy the corresponding action.

• It abstracts from actual Intent Sets that square measure outlined by the Intent suppliers. just in case the intent supplier is a dead ringer for the platform supplier, they'll match.

5.2.7 Dialog X

The Dialog X square measure able to handle practicality which will be superimposed to the capabilities of the Dialog Manager through their associated Intent Set X. Dialog X extends the Core Dialogs and add practicality by custom Dialogs. The Dialog X's should server completely different functions in a very sense that they're distinctive for an explicit task. E.g. , solely one flight reservation dialog could exists at a time. They have a same characteristics as a Dialog .

5.2.8 Intent Set X

An Intent Set X could be a special Intent Set that identifies tasks which will be dead among the associated Dialog X .

5.2.9 Dialog register

The Dialog register manages all out there Dialogs with their associated Intent Sets.

Dialogs and their Intent Sets will be superimposed or removed PRN.

The Dialog written account might give notice the   Dialog Management if dialogs are additional or removed.

5.3 APIs/Data Layer

5.3.1 Supplier choice Service

A service that gives access to any or all well-known IPA suppliers. This service additionally maps the IPA Intent Sets to the intent sets in Dialog Layer. It has the subsequent characteristics

• The supplier choice Service receives input as text strings and come back result as intents with all recognized entities from all IPA suppliers that area unit ready to reply to the user input at the side of associated entities.

• In case the supplier choice Service is named with a preselected IPA suppliers solely this one are going to be used

5.3.2 Accounts/Authentication

A written account that is aware of a way to access the well-known IPA supplier, i.e.that area unit offered and credentials to access them. Storing credentials should meet security and trust issues that area unit expected from such a personalised service.

5.3.3 Core NLU

• The Core NLU is in a position to handle basic practicality via Core Intent Sets to modify interaction with the user the least bit.

• The Core NLU might build use of the Core knowledge supplier to access native or internal knowledge or access internal services.

5.3.4 Core knowledge supplier

A generic knowledge supplier to assist the Core NLU decisive the intent.

5.3.5 IPA supplier X

A supplier of associate degree IPA service, like

• Google Assistant

• Amazon Alexa

• Microsoft Cortona

• SoundHound

5.3.5.1 Supplier NLU

An NLU element that's ready to extract that means as intents associate degreed associated entities from an auditory communication as text strings for IPA supplier X. it's the subsequent characteristics

• The supplier NLU might build use {of knowledge| of information} supplier to access native or internal data or access external services.

• The supplier NLU might build use of the data Graph to derive that means

5.3.5.2 Supplier Intent Set

An Intent Set that may be came by the supplier NLU to handle the capabilities of IPA supplier X.

5.3 .5.3 knowledge supplier

A data supplier to assist the supplier NLU in decisive the intent.

The information supplier provides access to o Local knowledge or External knowledge or External services.

5.3.5.4 Data Graph

A data graph to reason concerning the detected input from the supplier NLU and knowledge supplier to come back up with some a lot of meaningful results.

6. CONCLUSION

The findings from this paper provides insight on usage of virtual personal voice assistant that may be useful for lecturers in addition as in company world. Voice assistant area unit a lot of reliable than human personal assistant as a result of it reminds you vital things on correct things or location as per our want in any language.

7. ACKNOWLEDGMENT

REFERENCES

[1] D O'SHAUGHNESSY, SENIOR MEMBER, IEEE, "Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis" proceedings of THE IEEE,VOL.91,NO.9, Gregorian calendar month 2003

[2]Nil Goksel-Canbek Mehmet Emin Mutlu,"On the track of Artificial Intelligence: Learning with intelligent Personal Assistant" International Journal of Human Sciences.

[3] VINAY SAGAR, KUSUMASM, "Home Automation Using net of Things", June-2015, IRJET, e-ISSN:2395-0056.

[4]"Speech recognition with flat direct models," IEEE Journal of elect Topics in Signal Processings, 2010