International Journal of Advanced Technology in Engineering and Science
Vol. No. 09, Issue No. 08, August 2021
www.ijates.com
ijates
ISSN 2348 - 7550

A.I Desktop Voice Assistant

Dharmendra kumar, Chandan kushwaha, Bittu kumar, Deepak

[1]*"U.G Student Department of Computer Science and Engineering IIMT College of engineering, Greater noida ,Dr.A.P.J Abdul kalam Technical University, U.P,India"*
"dharmendrakumar121121@gmail.com"

[2]*"U.G Student Department of Computer Science and Engineering IIMT College of engineering, Greater noida ,Dr.A.P.J Abdul kalam Technical University, U.P,India"*
"Chandankushwaha775@gmail.com"

[3]*"U.G Student Department of Computer Science and Engineering IIMT College of engineering, Greater noida ,Dr.A.P.J Abdul kalam Technical University, U.P,India"*
"bittukumar85212@gmail.com"

[4]*"U.G Student Department of Computer Science and Engineering IIMT College of engineering, Greater noida ,Dr.A.P.J Abdul kalam Technical University, U.P,India"*
"dy2044010@gmail.com"

1. Abstract

The Most well known utilization of iPhone is "SIRI" which makes a difference the end client to convey end client versatile with voice and it likewise reacts to the voice orders of the client. Same kind of utilization is additionally evolved by the Google that is "Google Voice Search" which is utilized for in Android Phones. Be that as it may, this Application for the most part works with Internet Connections. However, our Proposed System has ability to work with and without Web Connectivity. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the client contribution to type of voice or text and cycle it and returns the yield in different structures like activity to be performed or the output is directed to the end client. What's more, this proposed framework can change the method of cooperations between end client and the cell phones. The framework is being planned in such a manner that every one of the administrations given by the cell phones are open by the end client on the client's voice orders.

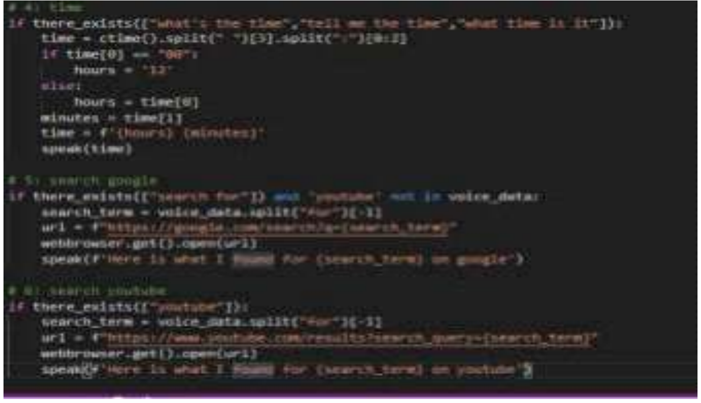
Keywords: SIRI, Google Voice Search, Mobile Device, Internet

2.Introduction

These days the Mobile Technology is by and large extremely well known for the User Experience, since it is not difficult to get to the applications and administrations from anyplace of your Geolocation. Android, Apple, Windows, Blackberry, and so forth are different acclaimed and regularly utilized Mobile Operating Frameworks. Every one of the Operating Systems gives a lot of applications and administrations for clients. For an occurrence, the Contacts Applications is utilized to store the contact subtleties of the client's contact and furthermore helps client to associate a call or send a SMS to other individual utilizing the substance put away in this application. We can get comparable sorts of application from one side of the planet to the other through Apple Store, Play Store, and so on This highlights brings forth different sorts of sensors or functionalities to be carried out in the cell phones. The Most well known use of iPhone is "SIRI" which makes a difference the end client to convey end client to versatile with voice what's more, it additionally reacts to the voice orders of the client. Same sort of use is additionally evolved by the Google that is "Google Voice Search" which is utilized for in Android Phones. However, this Application for the most part works with Internet Connections. In any case, our Proposed System has capacity to work with and without Internet Connectivity. It's named as Personal Assistant with Voice Recognition Knowledge, which takes the client contribution to type of voice or text and interaction it and returns the yield in different structures like activity to be performed or the query output is directed to the end client.

3. LITERATURE SURVEY

A discourse synthesizer takes as info and produces a sound stream as yield. A discourse recognizer then again does inverse. It takes a sound stream as info and hence transforms it into text record. The voice is a sign of endless data. An immediate investigation and orchestrating the unpredictable voice signal is expected to a lot of data contained in the sign. Thusly the advanced sign cycles, for example, Feature Extraction and Feature Matching are acquainted with address the voice signal. In this task we straightforwardly use discourse motor which use Feature extraction method .Our plan to make an ever increasing number of functionalities which can assist human with aiding their every day life and furthermore decreases their endeavors. Plan of a smaller enormous jargon discourse acknowledgment framework that can run effectively on any PC gadgets, precisely and with low inactivity. Discourse



```
# 4: time
if there_exists(["what's the time","tell me the time","what time is it"]):
    time = ctime().split(" ")[3].split(":")[0:2]
    if time[0] == "00":
        hours = '12'
    else:
        hours = time[0]
    minutes = time[1]
    time = f'{hours} {minutes}'
    speak(time)

# 5: search google
if there_exists(["search for"]) and 'youtube' not in voice_data:
    search_term = voice_data.split("for")[-1]
    url = f"https://google.com/search?q={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on google')

# 6: search youtube
if there_exists(["youtube"]):
    search_term = voice_data.split("for")[-1]
    url = f"https://www.youtube.com/results?search_query={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on youtube')
```

```
[illegible]
```

4. METHODOLOGY

1. **System Architecture:**

The overall system design consists of following phases:

(a) Data collection in the form of speech.

(b) Voice analysis and conversion to text

(c) Execute Python script

(d) Generating speech from the processed text output

In first phase, the data is collected in the form of speech and stored as an input for the next phase for processing. In second phase, the input voice is continuously processed and converted to text using STT[1]. In next phase the converted text is analysed and processed using Python Script to identify the response to be taken against the command.Finally once the response is identified, output is generated from simple text to speech conversion using

2. Data flow sequence

a. Instate gadget: Initialize the gadget by calling its name.

b. Errand Manager: Conversion of Speech-to-Text and Text-toSpeech is performed by task director.

c. Administration Manager: Analysis of orders and coordinating them with web administration and applications.

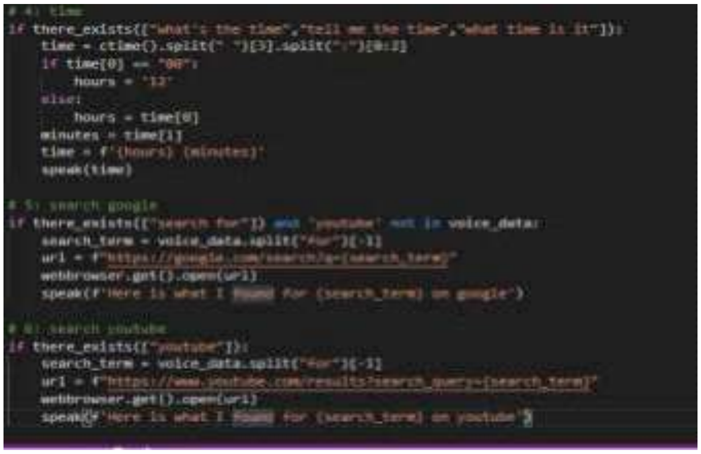
d. Execute Command: After discovering the counterpart for the given order, run the particular python content and give the yield.

5. Proposed System

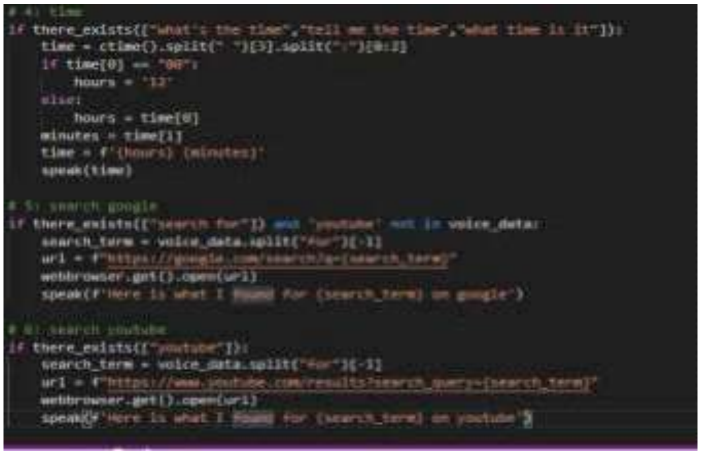
DVA has different parts of the administrations, yet the principle highlight of DVA is Voice Recognition Engine which has an capacity to work without web association for example Disconnected Voice Acknowledgment. DVA will accept an info information as a voice and given a yield information will finish the assignment. This Voice Recognizer works disconnected and performs different tasks according to the client orders and necessities. This is the first action that is opens at whatever point we start DVA and it straightforwardly can be opened by squeezing Power Button of the PC gadgets. This element is uncommonly plan for Blind People who wish to utilize the Laptop or PC gadgets be that as it may, can't associate this innovation. Additionally Native client who scarcely knows to on PC can undoubtedly open this application and utilizing voice orders in their nearby dialects according to require. DVA reacts to essential orders like, Open Applications, Close Applications, Connect Google Send Mail to particular individual , Add/Delete/Update File, Run any media record, Start different administrations like Hotspot, WiFi, Bluetooth, Music , Youtube and different Services from the particular Notification Panel. This can be performed on the voice orders of the end client without web network.

6. Technology used Python

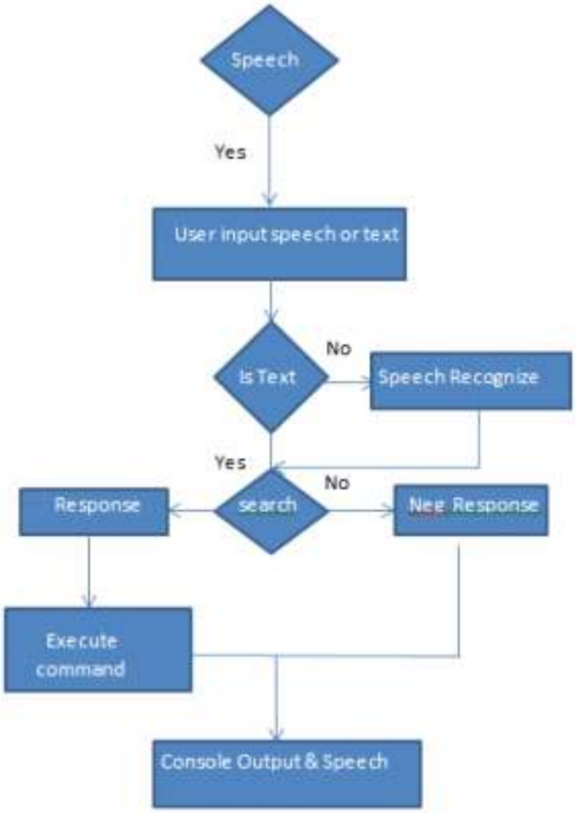
It is an overall significant programming language utilized in Dialog stream. It is utilized as a base for the most conspicuous Abased programming considering its flexibility, straightforwardness and longstanding standing. Python is a translator, undeniable level, broadly useful programming language.



```
# 4: time
if there_exists(["what's the time","tell me the time","what time is it"]):
    time = ctime().split(" ")[3].split(":")[0:2]
    if time[0] == "00":
        hours = '12'
    else:
        hours = time[0]
    minutes = time[1]
    time = f'{hours} {minutes}'
    speak(time)

# 5: search google
if there_exists(["search for"]) and 'youtube' not in voice_data:
    search_term = voice_data.split("for")[-1]
    url = f"https://google.com/search?q={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on google')

# 6: search youtube
if there_exists(["youtube"]):
    search_term = voice_data.split("for")[-1]
    url = f"https://www.youtube.com/results?search_query={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on youtube')
```



```
# 4: time
if there_exists(["what's the time","tell me the time","what time is it"]):
    time = ctime().split(" ")[3].split(":")[0:2]
    if time[0] == "00":
        hours = '12'
    else:
        hours = time[0]
    minutes = time[1]
    time = f'{hours} {minutes}'
    speak(time)

# 5: search google
if there_exists(["search for"]) and 'youtube' not in voice_data:
    search_term = voice_data.split("for")[-1]
    url = f"https://google.com/search?q={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on google')

# 6: search youtube
if there_exists(["youtube"]):
    search_term = voice_data.split("for")[-1]
    url = f"https://www.youtube.com/results?search_query={search_term}"
    webbrowser.get().open(url)
    speak(f'Here is what I found for {search_term} on youtube')
```



7.Conclusion

DVA is Designed to help Native and particularly for Blind people which deals with their Voice Commands. DVA likewise has the ability of perceiving the voice orders without web association. DVA has different functionalities of Laptop gadgets like organization association and overseeing different applications on the voice orders. Contains key highlights like Voice Pattern Detection, Keyword Learning, and so on which accommodating for end client to utilize different functionalities and administrations of the PC gadgets. Consequently, DVA is language hindrance autonomous which effectively reacts to client's voice orders quicker than the Online Voice Search applications. It is not difficult to utilize which it can accept a contribution as voice and satisfy the necessity of client.

8. Acknowledgement

We are truly pleased to present the Paper on "Work area Voice Assistant". We might want to communicate generously appreciation to Prof. Shabana Pathan who at extremely discrete advance in aiding us for readiness of this Paper and furthermore contributed his significant direction to help and address each issue that confronted. Additionally, we might want to communicate our genuine respects to our families for continually being there when we required them the most. With all regard and appreciation, we owe our prosperity to the scholars of reference papers that are alluded by us in finish of this administrative work action which will be valuable in introducing our review paper.

9. REFERENCES

[1] "Accurate and compact large vocabulary speech recognition on mobile devices," in INTERSPEECH. 2013, pp. 662– 665, ISCA.

[2] "Study of Voice Controlled Personal Assistant Device" by IJCTT Volume 42 Number 1 – December 2016

[3] "Personal Assistant with Voice Recognition Intelligence" ISSN 0974-3154 Volume 10, Number 1 (2017).

[4] "Language modeling for what-with Where on GOOG411, " Proc. of INTERSPEECH, 2009.

[5] "DFirect construction of compact context-dependency transducers from data," Proc. of INTERSPEECH, pp. 218– 221, 2010.

[6] "Automatic question generation for decision tree based state tying," Proc. of ICASSP, 1998.

[7] "Speech recognition with flat direct models, IEEE Journal of Selected Topics in Signal Processing, 2010.

[8] "Analysis and Comparison of Intelligent Personal Assistants" Oktay Bahcecia, aoktayb@kth.se.