INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

Offline speech recognition on android device based on supervised learning

Reeta Bandhu
reetakalyanpuri@gmail.com
Sinhgad Institute of Technology, Lonavala, Maharashtra

Nikhil Kumar Singh
nksingh121196@gmail.com
Sinhgad Institute of Technology, Lonavala, Maharashtra

Betawar Shashank Sanjay
shashankbetawar24@gmail.com
Sinhgad Institute of Technology, Lonavala, Maharashtra

ABSTRACT

Offline Android Smartphone Assistant is a virtual personal assistant which is implemented to perform basic functionalities of the smartphone in offline mode with speech commands. The task includes open apps, toggle Wi-Fi, toggle Bluetooth, call a person, send messages, control brightness, open torch, etc. The key functionality of the Offline Android Smartphone Assistant is the capability of performing each and every task without the use of the internet. The users can interact with the assistant through natural voice commands. Many companies had established several virtual assistants like Google Assistant, Apple's Siri, Amazon Alexa, etc. However, the virtual assistants established by several companies do not provide many offline functionalities. In this proposal, we provide easy to access every system functionality without using the internet using Speech Recognition. This application will help people in areas where internet connectivity is hardly possible. The application also provides the functionality of a screen overlay. The application uses Natural Language Processing for processing the voice commands and performing the required tasks.

Keywords— *Natural language processing, Speech recognition, Offline application, Voice commands, a Screen overlay*

1. INTRODUCTION

We have been using various kinds of a virtual personal assistant in our Android devices for a long time. Also, these virtual devices are evolving rapidly with their working system and efficiency of the system. But as we may have seen that almost all such personal assistant requires internet connectivity for their usage. These systems have a huge size of datasets due to which they cannot be stored offline, hence it requires online storage like cloud services. Hence, the application is implemented to provide all the system i.e. software as well as hardware functionalities accessible by the user through voice commands and without the use of internet connectivity.

This application uses the technology known as Natural Language Processing for processing the natural voice input given by the user to the system. The tasks features are extracted from the voice input and accordingly it is performed by the virtual assistant.

The application stores the datasets offline and works without the use of internet connectivity. It is seen that many smartphone users face problems to deal with system settings like toggling mobile data, turning on/off torch or finding an app, etc. This application would make users to easily operate any android device without any hassle just by giving voice commands as input. Also, it would be a great help in some cases of disabilities as the user would be able to operate his/her device just by giving voice commands to it.

1.1 Structure of offline android smartphone assistant

Many virtual assistants are implemented until now by several companies. According to research, in the future, almost every user will be using 'voice' for interacting with the machines. The aim of the implementation of the system is to provide ease to the user for operating the android devices and dealing with device settings without knowing the exact details about the system settings and without using the internet.

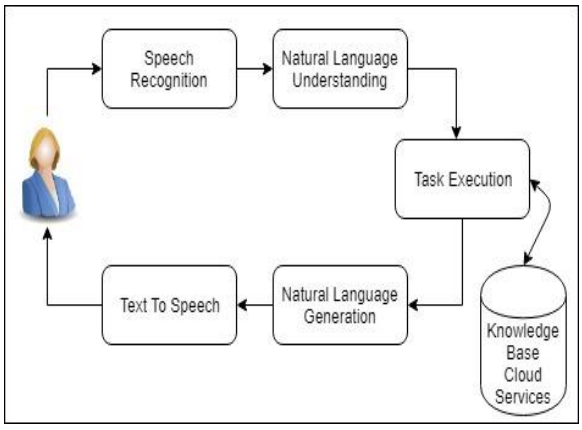

Fig. 1: General system architecture

The technology used in the implementation of the application is Natural Language Processing. By using this technology, the application takes input from the user in the form of natural voice commands. The voice command is then preprocessed. The preprocessed speech is used to get the exact meaning of the voice command and the relevant feature is extracted from the

command, for example, turn on Wi-Fi, open What Sapp, etc. The task is performed accordingly and informed to the user relatively. The application also provides screen overlay functionality which allows the application to run in the background over any running application. This feature provides the user to interact with the application at any point without bothering about opening the application again and again.

2. LITERATURE SURVEY

[2] Speech To Text Conversion Using Android Platform: The paper provides the implementation of an application which is based on Speech Recognition Algorithm i.e. Hidden Markov Model (HMM). It uses Android for the implementation purpose. The functionalities such as Read Contacts, Send Messages and Receive Messages are provided by the system using speech recognition.

[3] Architecture for Automatic Generation of User Interaction Guides with Intelligent Assistant: This paper has proposed the architecture for the automatic generation of user interaction guide with an intelligent assistant. It provides information about how to make the user familiar with the intelligent assistant as quickly as possible. The technologies like task network and Natural Language Generation/ TTS are used for the architecture. The task network is used to find the right sequence of the actions to execute the user's request. A plan recognition engine is also made to find the possible features which are relevant to the user's context. Finally, NLG/TTS is used to make the guide which is based on the plans which are recommended by the system. The limitation of the architecture is that it is limited to a few features. It does not deal with the handling of every system functionalities.

[4] Next Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa, and Google Home): The paper describes about the designing of Virtual Personal Assistants using Multi-model dialogue system which includes the techniques such as gesture recognition, image/video recognition, and speech recognition, the vast conventional knowledge base, and general knowledge base. The designing of the multi-model dialogue system with high accuracy is achieved by adding some components to the general dialogue system. These components include ASR Model, Gesture Model, Graph Model, Interaction Model, User Model, Input Model, Output Model, Interface Engine, Cloud Servers, and Knowledge Base.

[5] Smart Personal Assistant using Machine Learning: The paper proposes a machine learning approach for the learning mechanism of the personal assistant. It also describes better utilization of the batteries of the smartphone. The system uses a Support Vector Machine Algorithm which is a supervised machine learning algorithm. The algorithm is used for the classification of the features for the personal assistant. The application provides the functionality of user profile management which is done automatically without any human intervention. The profile management of the user includes the track of the locations frequently visited by the users, reminders, etc. The limitation of the application is that it only provides calling-based services and location-based services.

[6] Digital Life Assistant Using Automated Speech Recognition: This paper provides a brief description about automatic speech recognition, the technologies such as Hidden Markov Model (HMM) and Mel Spectrum Cestrum Coefficient (MFCC) are also discussed with their applications and future scope and aspects. HMM and MFCC are speech recognition algorithms. It also includes various speech recognition systems, such as Speaker Dependent Speech Recognition System and Speaker Independent Speech Recognition System. The paper provides a review of how these technologies actually work and how much progress has been made in the past few years.

3. OUR APPROACH

The Offline Android Smartphone Assistant uses the Natural Language Processing algorithm. The most important function of the system is to work offline i.e. without using internet connectivity. This application is helpful to the people with disabilities as it takes only voice commands as input. Also, it is helpful to the people who don't have internet connectivity and who don't know where the system settings provided in their phone are located.

The system basically works in 5 phases:

- Speech Input,
- Speech Preprocessing,
- Feature Extraction,
- Task Execution,
- Text To Speech.

These features are summarized below:
The user provides input in the form of voice commands. The speech of the speaker is received in the form of waves. The speech input does have background noise and room reverberation which is totally undesirable. Hence, to remove such noise, we process the speech input to get the user's context. Speech Preprocessing plays an important role in eliminating irrelevant sources of variations. It ultimately improves the accuracy of speech recognition. The preprocessed speech input is used to extract features from it. For example, open <App Name>, Toggle Bluetooth, Set Alarm, etc. After performing the desired task, the results in text format are converted into speech using Text To Speech algorithm and it is then informed to the user in the form of natural language speech output.

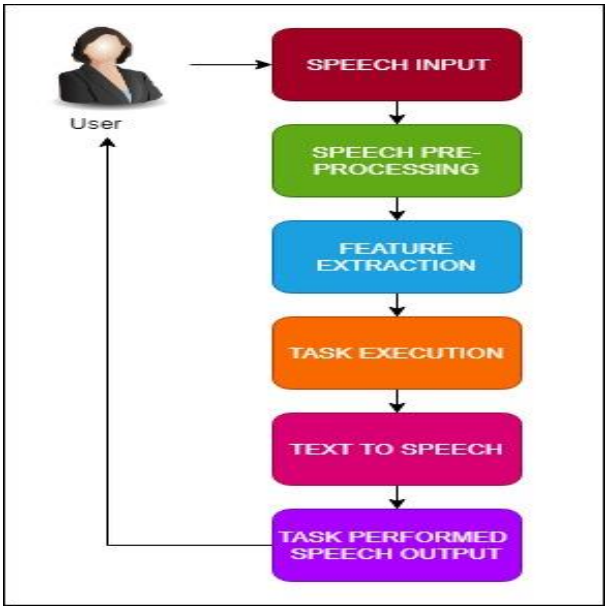

Fig. 2: Flow diagram of offline android smartphone assistant

3.1 Speech recognition

Speech Recognition in the application is achieved without using Google Server. We had used the Speech Recognizer API class available in Android Studio [1]. This class provides us with access to speech recognition services. Speech Recognizer API in Android System deals with supervised machine learning algorithm which includes Natural Language Understanding. Here, the machine is trained using labelled datasets. After that,

the machine is provided with a new set of data which is known as the testing dataset. The machine predicts the output of testing data set after analyzing training dataset.

To use Speech Recognizer API class, desired permission for recording the audio i.e. permission. RECORD_AUDIO should be included in Android. Manifest file [2]. Speech Recognize .start Listening () method allow us to access the microphone for the user's command. A Recognition Listener interface is used to receive notifications from Speech Recognizer when the recognition related events occur. The Recognition Listener also deals with the call-backs. Speech Recognizer class takes the voice command as an input and convert the speech into text using Speech to Text Algorithm. Speech Recognizer deals with speech pre-processing as well. The Pre-Processing of speech includes filtering the voice command and removing the noise that is present in the command. For instance, if a person says "hello" very quickly and another person might say "hhheeeellllllooooooo" very slowly then both the sound files are recognized as "hello" by the Speech Recognizer.

3.2 Action

In this phase, the text generated by the Speech Recognizer class is used to perform the request made by the user through voice command. The feature extraction process is done here. In the feature extraction process, the data in text format is used to get the user's context. Here, we compare an intent with the text and if that particular intent is found in the text, then the desired task is executed.

For example, the user provides input as "What is the time". Here, "time" is the intent/feature which is to be extracted from the text. After getting time, the user is responded with the current time in the form of speech.

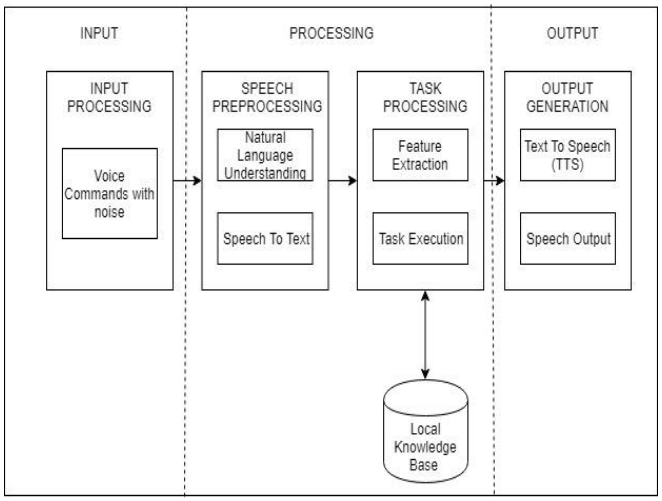

Fig. 3: Generalized architecture of offline android smartphone assistant

3.3 Response

The application provides output in a speech format. After executing the task which is requested by the user, the user is informed about the completion of the task in a speech format. Here, the android studio provides a library android.speech.tts which converts the Text into Speech. The library android.speech.its includes TextToSpeech class which synthesizes speech from the text for immediate playback.

The TextToSpeech class has a method known as speak() which allow us to convert any text to be spoken by the machine. For example, speak("Bluetooth is turned on") will make the machine speak "Bluetooth is turned on" which is very useful for proper interaction between user and machine.

3.4 Features implemented by the application

- Toggle wi-fi.
- Toggle Bluetooth.
- Manipulate Brightness.
- Screen Rotation.
- Call Emergency numbers such as- Police, Fire Alarm, Women Helpline numbers.
- Screen Overlay – the application can be used over any other running application.
- Alam Remainders.
- Changing Profile (Silent/ Vibrate/ Ring).

4. FUTURE SCOPE

The application can be further developed for various languages like Hindi, Russian, French, etc. The functionalities such as Calling a person, Messaging, Flight Mode, etc may be included, which will be helpful to the people who don't have internet connectivity, also to the people with disabilities.

5. CONCLUSION

On studying the various applications which are already available in the market, we have seen that almost every application work with internet connectivity. Also, some of them are not very efficient with speech recognition. The Offline Android Smartphone Assistant which is implemented provides the usage of the application without internet connectivity. The user can use the application to deal with any of the system settings either software or hardware by just giving voice commands to the application.

The application also provides screen overlay functionality, which allows the application to run over any other running application. The user can access the application to give voice commands without bothering to open it again and again just by clicking on the mic button on the screen. The application would be of great use in rural areas where internet connectivity is barely possible. Also, it would help the person with disabilities to operate phone as the user have to provide voice command only.

6. REFERENCES

[1] https://developer.android.com/reference/
[2] B. Raghavendhar Reddy and E. Mahender, "Speech To Text Conversion using Android Platform", International Journal of Engineering Research and Application, Jan-Feb 2013.
[3] Inchul Hwang, Jinhe Jung, Jaedeok Kim, Youngbin Shin, and Jeong-Su Seol, "Architecture for Automatic Generation of User Interaction Guides with Intelligent Assistant", 2017 31st International Conference on Advanced Information Networking and Applications Workshops,2017 IEEE.
[4] Veton Kepuska and Gamal Bohouta, "Next-Generation of Virtual Personal Assistants(Microsoft Cortana, Apple Siri, Amazon Alexa, and Google Home)", 2018 IEEE.
[5] Prajyot Mane, Shubham Sonone, Nachiket Gaikwad, and Prof. Jyoti Ramteke, "Smart Personal Assistant using Machine Learning", International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017), 2017 IEEE.
[6] Seema Rawat, Parv Gupta, and Praveen Kumar, "Digital Life Assistant using Automated Speech Recognition", International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH14) 28 & 29 November-2014.

