VIRTUAL PERSONAL ASSISTANT USING ARTIFICIAL INTELLIGENCE

1Meenakshi Garg  
2Kiran Bala  
3Sakshi Sharma  

123Assistant Professor  
1,2,3Department of Computer Science Engineering, Chandigarh Group of College, Jhanjeri, Mohali, India

Abstract—An intelligent virtual assistant (IVA) or intelligent personal assistant (IPA) may be a software agent which will perform tasks or services for private supported commands or questions. Sometimes the term "chatbot" is employed to ask virtual assistants generally or specifically accessed by online chat. In some cases, online chat programs are exclusively for entertainment purposes. Some virtual assistants are ready to interpret human speech and respond via synthesized voices. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks like email, to-do lists, and calendars with verbal commands.

Keywords- VPA, NLP, Speech to text, Text Analyzing, Artificial Intelligence.

INTRODUCTION

People no longer rely on other people for assistance or services. Humans no longer need to ask for assistance from others since they can rely on a far more effective and dependable equipment to take care of their daily needs as a result of the world becoming increasingly digital. The use of computers, mobile devices, laptops, and other similar devices has permeated every aspect of our lives. These devices can run both simple and complicated programmed, which helps to cut down on boring labor and manpower waste.

Virtual personal assistants have practically become a basic requirement in all electronic gadgets to resolve the problem quickly. To implement this, speech recognition became the new integration into the VPA. Virtual Personal Assistant has almost become a basic necessity in all electronic devices so as to execute the required problems easily. More than just being a bot, VPA can make life easier for the user in various ways. Speech recognition is one of the relatively new integrations into the VPA. But, though its moderately efficient, it is not very helpful and are not used by the user due to its high amount of error. Though the error percentage of the upcoming VPAs is around 5 percent, it still isn’t quite up to the mark to where it becomes a basic part of the user’s life. Thus, the projects aim is to build a VPA with speech recognition which has a very minimal error percentage.
Voice recognition is a complex process using advanced concepts like neural networks and machine learning. The auditory input is processed and a neural network with vectors for each letter and syllable is created. This is called the data set. When a person speaks the device compares it to this vector and the different syllables are pulled out with which it has the highest correspondence. The fact that the car has evolved into mobile office and safety became a measure concern for it. According to the Statista, there will be over 8 billion digital voice assistants in use worldwide by 2024, roughly equals to the world population. It is estimated that it will be worth several billions by 2007. While indirect revenues for the carriers will be several folds. A few companies have started offering converging products in the VPA direction, e.g., Conita, WildFire, VoxSurf, VoiceGeneie, and VoiceTel, and Mitel Networks, though own or two methods will provide the solutions for mobile carrier environment. At last, it provides hands-free, eyes-free access to the web anywhere, anytime from any phone. Thus, the project’s goal is to create a very basic activated assistant (VPA) using speech recognition and plays the songs from the You Tube.

**EXISTING MODEL**

The majority of current efforts have simply used neural networks for speech recognition. Despite having a decent level of accuracy, these techniques are neither efficient nor practical to be of any meaningful value. They employ a few simple strategies, including:

1. **Context-aware computing**:

   This Context-aware system has the ability to sense their physical environment and adapt themselves accordingly as it suggests from its name. The words speak by the people in different accent, that can also be recognized using this method. It also automatically deduces the words which are spoken twice a time.

2. **MFCC**:

   MFCC stands for the Mel-Frequency Cepstral Coefficients. It works on the collection of these coefficients. It adds up to the transient power range of a sound. These can be utilized to detect varieties in sound to perceive the different factors expected for voice acknowledgment.

3. **NLP**

   The study of how human and computer languages interact is the focus of the Artificial Intelligence subfield known as Natural Language Programming. It focuses primarily on how to program computers to process the vast amount of data on natural languages. This idea is used to train the computer to recognize spoken words and familiarize itself with the various words in a particular language.

**TECHNIQUE AND METHODOLOGIES USED IN PROPOSED MODEL**

This project's system design demonstrates how control moves throughout the system. When user demand for a song then it directly plays from the You Tube. The following is the architecture diagram. The majority of current efforts have simply used neural networks for speech recognition. Despite having a decent level of accuracy, these techniques are neither efficient nor practical to be of any meaningful value. They employ a few simple strategies, including:
SPEECH-TO-TEXT (STT) AND TEXT-TO-SPEECH (TTS)

Speech-to-text – It allow the applications to translate the spoken words into digital signals. At the point when you talk, you make a progression of vibrations. The software converts them into digital signals with an analog-to-digital converter (ACD), extracts sounds, segments them, and matches them to existing phonemes. The smallest unit of a language that is capable of distinguishing the sound shells of various words are called phonemes. The system creates a text version of what you said by comparing these phonemes to individual words and phrases using intricate mathematical models.

Text-to-speech - It concept is entirely opposite too previous one. This technology translate text into voice output learning. The system must go through three steps to convert text to voice. First, the system needs to convert text to words, then perform phonetic transcription and then convert transcription to speech.

Speech-to-text (STT) and Text-to-speech (TTS) are used in virtual assistant technology to ensure smooth and efficient communication between users and applications. To turn a basic voice assistant with static commands into a proper AI assistant, you also need to give the program the ability to interpret user requests with intelligent tagging and heuristics.

![Diagram](Fig 3)
(The working of Virtual Assistant uses following principles)

- **Natural Language Processing**: Natural Language Processing (NLP) refers to AI method of communicating with an intelligent system using a natural language such as English. Processing of Natural Language is required when you want an intelligent system like robot to perform as per your instructions, when you want to hear decisions from a dialogue based clinical expert system, etc.

![Diagram](Fig 4)
(Five Steps in Natural Language Processing)

Automatic Speech Recognition: To understand command according to user’s input.
Artificial Intelligence is the concept to learn from the user and store all of their behavior and relationships information. The capacity of a system to calculate, reason, perceive relationships and analogies, learn from experience, store and retrieve information from memory, solve problems, comprehend complex ideas, fluently use natural language, classify, generalize, and adapt to new circumstances.

Inter Process Communication: To get important information from other software applications.

CONCLUSION & FUTURE SCOPE

A brand-new mobile service is described in the paper. For mobile professionals, the Virtual Personal Assistance offers an intelligent computer secretarial service. The new service is based on the convergence of internet, mobile, and speech recognition technology. The VPA provides a single point of communication for all of the user’s messages, contacts, schedule, and information sources, reducing interruptions and enhancing time utilization. The paper also suggests a decision structure for handling appointment and meeting request requests as well as call screening. The framework at first targets legal counselors, specialists, deals work force, little workplaces, upkeep teams, and so forth. However, millions of additional users are anticipated to adopt it as a standard feature. It gets around many of the problems with the other solutions. It is mostly designed to make a VPA that works much better so that they can be used in more everyday situations. However, the system has limitations of its own. Despite its high efficiency, the time it takes to complete each task may be longer than that of other VPAs, and the complexity of the algorithms and concepts makes it difficult to modify in the future.

REFERENCES

Annual International Conference on Mobile Computing and Networking (pp. 343-355). ACM.

