



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: VI Month of publication: June 2020

DOI: <http://doi.org/10.22214/ijraset.2020.6092>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Artificial Intelligence for Speech Recognition

Dr. Mamatha G¹, Monika Raj², Swathi S³, Vinay G⁴

¹Head of the Department, ^{2,3,4}Student, Information Science Engineering, Nagarjuna College of Engineering and Technology, Karnataka, Bangalore, India

Abstract: Artificial Intelligence (AI) for speech recognition talks about the study and design of intelligent agents & also used to describe a property of machines or programs. It makes machines smarter and more useful and is less expensive than natural intelligence. Speech recognition converts spoken words to machine-readable input. It is also called Voice Recognition. The paper deals with various aspects of Speech recognition. Speech recognition includes- Voice dialling, Content-based spoken audio search, Speech-to-text processing, Performance of speech recognition systems. The paper states that speech recognition is usually specified in terms of accuracy and speed. Accuracy may be measured in terms of performance accuracy which is usually rated with word error rate, whereas speed is measured with the real time factor. This paper presents the Speech Recognition in Artificial intelligence systems and it is important to consider the environment in which the speech recognition system has to work.

I. INTRODUCTION

Artificial Intelligence (AI) involves two basic ideas. First, it involves studying the thought processes of human beings. Second, it deals with representing those processes via machines (like computers, robots, etc).

Natural language processing (NLP) refers to artificial intelligence methods of communicating with a computer in a natural language like English. The main objective of an NLP program is to understand input and initiate action. The input words are scanned and matched against internally stored known words. Identification of a keyword causes some action to be taken. In this way, one can communicate with the computer in one's language. No special commands or computer language are required. There is no need to enter programs in a special language for creating software. Voice XML takes speech recognition even further. Instead of talking to your computer, you're essentially talking to a web site, and you're doing this over the phone. OK, you say, well, what exactly is speech recognition? Simply put, it is the process of converting spoken input to text. Speech recognition is thus sometimes referred to as speech-to-text. Speech recognition allows you to provide input to an application with your voice. Just like clicking with your mouse, typing on your keyboard, or pressing a key on the phone keypad provides input to an application; speech recognition allows you to provide input by talking. In the desktop world, you need a microphone to be able to do this. In the Voice XML world, all you need is a telephone. When you dial the telephone number of a big company, you are likely to hear the sonorous voice of a cultured lady who responds to your call with great courtesy saying 'welcome to company X. Please give me the extension number you want'. You pronounce the extension number, your name, and the name of the person you want to contact. If the called person accepts the call, the connection is given quickly. This is artificial intelligence where an automatic call-handling system is used without employing any telephone operator. AI is behaviour of a machine, which, if performed by a human being, would be called intelligence. It makes machines smarter and more useful, and is less expensive than natural intelligence. The concepts used in AI include the principles outlined by man machine interfacing (MMI) which allows the creation of machines that are more usable for humans. Speech and gestures are the natural means of communication used by humans to interact with each other.

II. SPEECH RECOGNITION

Speech Recognition is a process of speech signals into a sequence of words. In 1990's Speech recognition reached a practical level with a limited satisfaction. Speech recognition is the one of the main benefit. User can concentrate on manual and observation operations by using the voice input commands. Computer speech recognition is the quite convenient, but most of the users using mouse and keyboard for the most convenient. Speech Recognition System is usually specified in terms of accuracy and speed. Accuracy may be measured in terms of performance accuracy which is usually rated with word error rate, whereas speed is measured with real time factor. Dictation machines can achieve very high performance in controlled conditions and require only a short period of training.

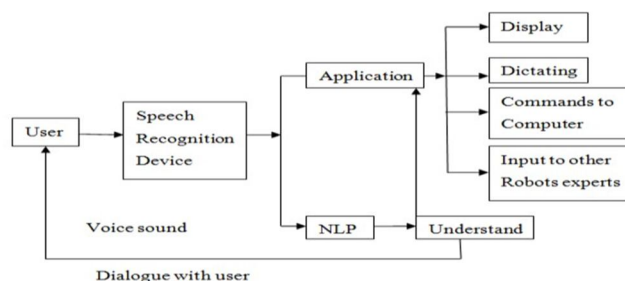
Optimal conditions usually assume that users:

- 1) Have speech characteristics which match the training data
- 2) Can achieve proper speaker adaption
- 3) Work in clean and no noise environment

III. SPEECH RECOGNITION PROCESS

After the training process, the user's spoken words will produce text; the accuracy of this will improve with further dictation and conscientious use of the correction procedure. With a well-trained system, around 95% of the words spoken could be correctly interpreted. The system can be trained to identify certain words and phrases and examine the user's. However, there are many other factors that need to be considered in order to achieve a high recognition rate. There is no doubt that the software works and can liberate many learners, but the process can be far more time consuming than first time users may appreciate and the results can often be poor. This can be very demotivating, and many users give up at this stage. Quality support from someone who is able to show the user the most effective ways of using the software is essential.

IV. SPEECH RECOGNITION SYSTEM



V. USES AND APPLICATIONS

A. Dictation

Dictation is the most common use for ASR systems today. This includes medical transcriptions, legal and business dictation, as well as general word processing. In some cases special vocabularies are used to increase the accuracy of the system.

B. Command and Control

ASR systems that are designed to perform functions and actions on the system are defined as Command and Control systems. Utterances like "Open Netscape" and "Start a new xterm" will do just that.

C. Telephony

Some PBX/Voice Mail systems allow callers to speak commands instead of pressing buttons to send specific tones.

D. Wearables

Because inputs are limited for wearable devices, speaking is a natural possibility.

E. Medical/Disabilities

Many people have difficulty typing due to physical limitations such as repetitive strain injuries (RSI), muscular dystrophy, and many others. For example, people with difficulty hearing could use a system connected to their telephone to convert the caller's speech to text.

F. Embedded Applications

Some newer cellular phones include C&C speech recognition that allow utterances such as "Call Home". This could be a major factor in the future of ASR and Linux. Why can't I talk to my television yet?

VI. CONCLUSION

Speech recognition has definite potential for reducing pilot workload, but this potential was not realized consistently. Achievement of very high recognition accuracy (95% or more) was the most critical factor for making the speech recognition system useful – with lower recognition rates, pilots would not use the system. More natural vocabulary and grammar, and shorter training times would be useful, but only if very high recognition rates could be maintained.



REFERENCES

- [1] R.Sathya, M.Pavithra, G.Girubaa, "Artificial Intelligence for Speech Recognition", International Journal of Computer Science & Engineering Technology
- [2] Manish G. Gohil "Artificial Intelligence for Speech Recognition", Volume-1, Issue-2, Sept 2014, International Multidisciplinary Research Journal.
- [3] <https://www.scribd.com/doc/30102999/Artificial-Intelligence-for-Speech-Recognition>
- [4] McCarthy, J. (1979) Ascribing mental qualities to machines. In: Philosophical perspectives in artificial intelligence, ed. M. Ringle. Atlantic Highlands, N.J.: Humanities Press.
- [5] Haugeland, J. (ED). (1985), Artificial intelligence: The very Idea, Massachusetts Institute of Technology, Massachusetts : MIT Press.
- [6] Kurzweil, R. (1990). The age of Intelligent Machines, Massachusetts Institute of Technology, Massachusetts : MIT Press.
- [7] Charniak and Mc Dermoth, (1985). Introduction to Artificial Intelligence, USA: Addison – Wesley.
- [8] Nilson, N.J. (1998). Artificial Intelligence – A new synthesis. Morgan Koufmann.
- [9] <http://research.microsoft.com/enus/news/features/speechrecognition-082911.aspx>
- [10] <http://dl.acm.org/citation.cfm?id=1752355>
- [11] <http://www.creativecow.net/interstitial.php?url=http%3A%2F%2Fforums.creativecow.net%2Fthread%2F279%2F626&id=0>
- [12] www.ijscce.org/attachments/File/v2i5/E1054102512.pdf
- [13] http://en.wikipedia.org/wiki/Outline_of_artificial_intelligence
- [14] http://www.csd.cs.cmu.edu/research/areas/vis_speech_lang/
- [15] <https://www.seminaronly.com/electrical%20&%20electronics/AI%20for%20Speech%20Recognition.php>
- [16] <http://tldp.org/HOWTO/SpeechRecognition-HOWTO/introduction.html>



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)