AI Receptionist

Mr. Kevin Joy D'Souza[1], Ms. P Swati[2], Ms. Aishwarya Shetty[3], Ms. Simpa V S[4]
[1]Assistant Professor, Dept. of CSE, Yenepoya Institute of Technology, Moodbidri, India-574225
Students, Dept. of CSE, Yenepoya Institute of Technology, Moodbidri, India-574225

Abstract - Speech is one in all the necessary modes of communication. Speech also can act as associate degree interface to speak with completely different machines. Attributable to this reason, machine-driven speech recognition systems are being enforced in numerous areas. one in all the samples of such system is associate degree AI receptionist. This paper primarily focuses on building a system which might answers the user queries using Artificial Neural Network (ANN). The words and sentences are trained by using Natural Language Processing (NLP) and linguistics search is employed for pattern matching.

Key Words: Artificial Neural Network; Natural Language Processing

1. INTRODUCTION

Speech is one in all the foremost necessary mode of communication. From the childhood, individuals do communication with one another through speech. It also can be used an efficient mode to move with the machines. In recent few years, the necessity for machine-driven speech recognition systems is very inflated and it's being enforced in numerous areas. It additionally led to the evolution of microphone, mobile etc.

Even in the eighteenth-century individuals tried to create speech recognition systems. For instance, Von Kempelen developed a machine that is capable of speaking words and phrases. Attributable to the evolution of machine power, currently it's become doable to develop, take a look at and implement speech recognition systems, and additionally to own systems which might convert time period text into speech. Despite being an excellent progress during this field still there are a unit ton of issues, as a result of speech could be a terribly subjective development. Speech is greatly stricken by accents, articulation, pronunciation, roughness, emotion, gender, pitch, speed, volume, ground noise and echoes.

Speech Recognition or Automatic Speech Recognition (ASR) plays a very important role in human laptop interaction. Speech recognition uses the technology to convert speech signals into the sequence of words. In theory, there ought to be the likelihood of recognition of speech directly from the digitized wave shape [1]. At present, speech recognition systems area unit capable of beneath standing of thousands of words under useful setting.

Speech signal provides two necessary sorts of information: (a) context of speech and (b) speaker's identity. Speaker recognition deals with the extraction of speaker's identity [2]. Speech recognition technology is employed in varied applications. it's already utilized in live subtitling on tv, as dictation tools in medical and legal community and for off-line speech-to-text conversion or note-taking systems [3]. It additionally has several applications like phone book help, automatic voice translation into foreign languages, spoken info querying for brand spanking new and inexperienced users and handy applications in field work, artificial intelligence and voice-based commands [4].

2. BACKGROUND STUDY

2.1 Speech Recognition

Speech recognition system area unit the programs that transforms the sound waves created by an individual into a graphical kind like text. the subsequent figure shows the steps concerned in speech recognition [5]:

Fig -1: Speech Recognition Method

Speech is that the vocalized style of human interactions. during this step, the speech of the speaker is received in wave shape. We are able to have ground noise or area reverberation in conjunction with the speech signal that is totally undesirable. To resolve this drawback, Speech *Pre-processing* is employed, within which moot sources area unit eliminated from the input speech signal. This step involves filtering, smoothing, framing, windowing, reverberation cancelling and echo removing etc.

Feature extraction is that the next step. The speech varies from person-to-person. this can be because of the very fact that each person has completely different characteristics embedded in auditory communication. In theory, risk ought to be there to acknowledge speech from the digitized wave shape. However, because of the big variation in speech signal, there arise a necessity to perform some feature extraction to scale back that variations.

Speech classification is that the step within which involve advanced mathematical functions and that they eliminate hidden data from the input processed signal.

Recognition is that the final step within which the words enclosed within the speech is recognized and data is deduced.

2.2 Artificial Neural Network

Artificial Neural Networks (ANN) are nothing however the crude electronic models supported body structure of brain. The human brain essentially learns from the experiences. it's an indisputable fact that some issues that area unit on the far side the scope of current computers is area unit simply soluble by energy economical packages. Such sort if brain modelling additionally provides a less technical path for the event of machine answer. ANN area unit laptop having their design modelled once the brain. They primarily involve many easy process units wired along in advanced communication network. Every easy process unit represents a true somatic cell that sends off a replacement signal or fires if it receives a robust signal from the opposite connected unit.

Artificial Neurons area unit the essential unit of Artificial Neural Network that simulates the four-basic operate of biological somatic cell. It's a mathematical relation planned as a model of natural somatic cell. The subsequent figure shows the essential artificial somatic cell.
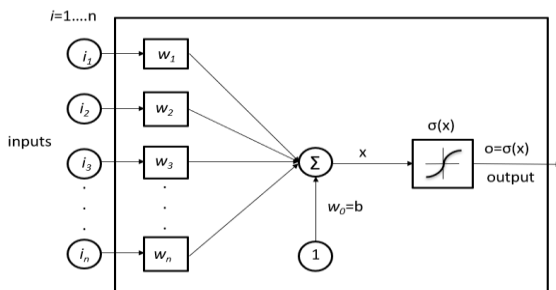

Fig -2: Artificial somatic cell

In the Figure-2, varied inputs area unit shown by the grapheme, i(n). Every of those inputs area unit increased by connecting weights w(n). Generally, this product is solely summed and fed to the transfer operate to get the output results. The applications like text recognition and speech recognition area unit needed to show these real-world inputs into separate values. These applications don't continually utilize networks composed of neurons that merely add, and thereby swish, inputs. Within the code packages, these somatic cells area unit referred to as process components and have more capabilities than the essential artificial neuron delineated higher than.

2.3 Natural Language Processing

Natural Language processing (NLP) deals with the interaction between humans and computers employing a language like English. natural language processing technology applies machine learning algorithms to text and speech.

The steps concerned in natural language processing are:

- Sentence Segmentation
- Word Tokenization
- Predicting elements of speech for every token
- Text Lemmatization
- Distinctive stop words
- Dependency parsing and finding noun phrases
- Named Entity Recognizer
- Conference Resolution

2.4 Parts of Speech Tagging

Semantic internet provides the users a temperature and reduces the wastage of time [6]. The projected work on linguistics search is accomplished by POS (Parts of Speech) tagging victimization language process. Victimization POS Tagging, the proposed linguistics search rule will perceive what the user question conveys and thus it provides additional relevant results to the user. Whenever a user question arises, it's given to Stanford programme for POS tagging. Supported POS Tagging, NN (Noun, Singular), NNS (Noun, Plural), NNP (Proper Noun, Singular), JJ (Adjective) and IN (Preposition) square measure extracted from user question so solely specific keywords are often chosen from the question that reduces search time.

3. IMPLEMENTATION

The enforced system includes tools that facilitate to alter the work of coaching associate instance of the system. The coaching of the system includes storing the queries and connected answers within the info as input. The coaching knowledge forms a network that consists of words, phrases etc.
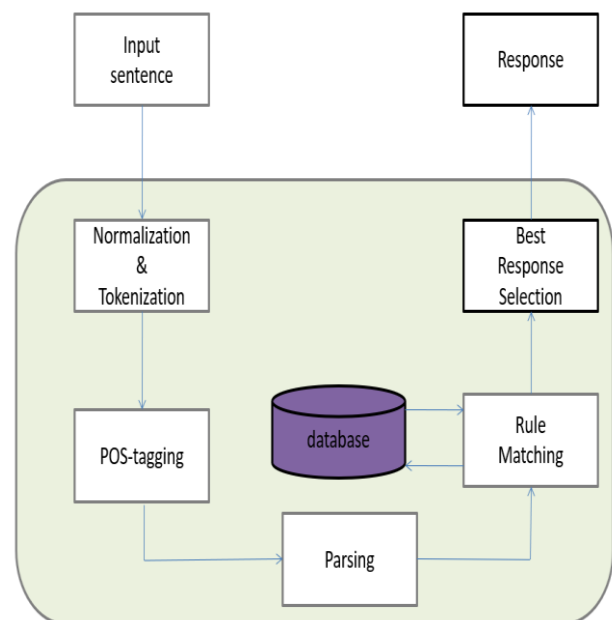

Fig -3: System Design

The flow of events of the system is as follows:

- First, once the user starts the system, he uses a mike to send out the input. Basically, what it will is that it takes sound input from the user and it's fed to the system to method it any.
- Then the input sentence is tokenized and POS tagging takes place wherever noun, verb, adjective etc. square measure known.
- The voice command system is made round the system of keywords wherever it searches the text for key words to match. And once key words square measure matched for a question within the database then it selects the relevant output.
- This output is displayed on the screen and conjointly it's transmitted via the mike.

- If the input question isn't found within the database, then the system can offer "Unknown" because the output.

4. RESULTS

When the user connects the mike and runs the system the system asks the user for speech input. After recognizing the speech input, the system can explore for the database for the user given question victimization natural language processing and produces the relevant output within the sort of each text and speech.
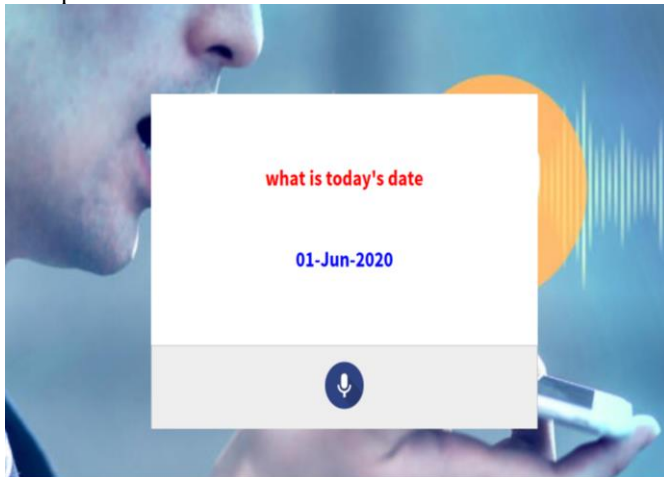

Fig -4: Speech Recognition and display of output

CONCLUSION

In this paper, we've got enforced an automatic speech recognition system which may offer answer to the queries asked by the user. The substitute Neural Network and machine learning is employed for implementing this method. The user can raise his or her queries to the system. The system takes this speech input and it'll extract the right keyword from the user question and can turn out the speech output. This method is often used as an automatic receptionist at the reception. So, during this approach we've got enforced an automatic speech recognition system.

REFERENCES

[1] Vimal Krishnan VR, Athulya Jayakumar, Babu Anto P, "Speech Recognition of Isolated Malyalam Words Using Wavelet Feature and Artificial Neural Networks", 4th IEEE International Symposium on Electronic Design, Test and Application, 2008.

[2] Shivam sharma, "Speech Recognition with Hidden Markov Model", in proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP'76) (2015), Pennsylvania, Vol. 24, No. 3, pp.201-212.

[3] Ms. Sneha K. Upadhyay, Mr. Vijay N. Chavda, "Intelligent system based on speech recognition with capability of self-learning", International Journal for Technological Research in Engineering ISSN (Online): 2347 - 4718 Volume 1, Issue 9, May-2014.

[4] Schultz, Tanja, Ngoc Thang Vu, and Tim Schlippe. "GlobalPhone: A multilingual text & speech database in 20 languages." Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2013.

[5] Bhushan C. Kamble, "Speech Recognition Using Artificial Neural Network –A Review", Int'l Journal of Computing, Communications & Instrumentation Engg. (IJCCIE) Vol.3, Issue 1 (2016) ISSN 2349-1469 EISSN 2349-1477.

[6] Sudhakar Pandiarajan, V.M. Yazhmozhi and P. Praveen kumar, "Semantic Search Engine Using Natural Language Processing".