

Speech Based Online Food Ordering System

Omkar Kulkari,Sakshi Wasekar,Gauri Dangre,Hrishikesh Dongre,Pradnya kasture

^{1,2,3,4} Computer Engineering Student, RMD College of Engineering, Pune, India, Guided by Prof Pradnya Kasture

Abstract- Use of speech recognition in the modern era is in the boost as it is easy to use. Nowadays electronic devices are now trained to learn by themselves and are able to interact with the user using voice commands. So in this paper we are proposing to develop a food ordering system using the speech recognition. The main focus to develop the system is which can be easy to use all the age group

Keywords- Speech Recognition, Natural language processing (NLP), Database, Google API, Spacy

I. LITERATURE SURVE

Sr.No	Title	Year	Description
1	Bing Speech API.	2018	It converts spoken audio to text for intuitive interaction
2	Web Speech API,” Technical Report, KTH Royal Institute of Technology	2018	The focus lies on measuring the performance of Google's Speech Recognition
3	Nuance Recognizer	2017	In this paper Nuance Recognizer used for NLP it delivers seamless self service and the total cost savings enabled by nuance security suite..
4	Pro Microsoft Speech Server2007: Developing Speech Enabled Applications with .NET (Pro). Apress	2007	three Primary components are Speech SDK, Telephony, ASP
5	Speaker recognition using hidden Markov models, dynamic time warping and vector quantization	1995	methods are used to combine HMM trained for speech and noise,here 60 % accuracy

II. INTRODUCTION

In this article we are designing a food ordering system which is based on the speech recognition. In this food ordering system the audio is taken as the input. The user is who is ordering can give the order by voice commands as the input instead of the traditional way of typing or selecting from UI. To make the outcome hustle free and quick we are filtering the input audio in

such way that it only search for the keywords in the database and shows the desired outcome. The fundamental aspect of speech recognition is the translation of sound into text and commands. Speech recognition is the process by which computer maps an acoustic speech signal to some form of abstract meaning of the speech. This conversation of the speech is achieved by the help of Google API and the Natural language processing(NLP).

Types of speech recognition system based on utterances.

Isolated Words

Isolated word recognition system which recognizes single utterances i.e. single word. Isolated word recognition is suitable for situations where the user is required to give only one word response or commands, but it is very unnatural for multiple word inputs. It is simple and easiest for implementation because word boundaries are obvious and the words tend to be clearly pronounced which is the major advantage of this type. The drawback of this type is choosing different boundaries affects the results.

Connected Words

A connected words system is similar to isolated words, but it allows separate utterances to be “run-together” with a minimal pause between them. Utterance is the vocalization of a word or words that represent a single meaning to the computer.

Continuous Speech

Continuous speech recognition system allows users to speak almost naturally, while the computer determines its content.

Basically, it is computer dictation. In this closest words run together without pause or any other division between words. Continuous speech recognition system is difficult to develop because system consider huge pause as end of the sentence.

Spontaneous Speech

Spontaneous speech recognition system recognizes the natural speech. Spontaneous speech is natural that comes suddenly through mouth. An ASR system with spontaneous speech is able to handle a variety of natural speech features such as words being run together. Spontaneous speech may include mispronunciation, false-starts and non-words.

III. MOTIVATION

Nowadays every work around us can be done giving specific set of commands. Like Siri Developed by apple, Alexa Developed by Amazon. So looking at this amazing concept and the future of technology, so it is very enthusiastic to develop an application based on speech. This will be very useful for everyone as now they can provide order just by voice command. And it is very easy as people do not require to use UI. The NLP is done with help of the spacy. Spacy is not a Platform or a API its is open source library designed for developing NLP application

IV. PROJECT SCOPE

1. Food ordering applications.
2. Order food online using speech instead of using UI.
3. Process the instruction provided.
4. Place the food order specified.

V. PROPOSED SYSTEM

The system proposed by is developed using the spacy for Natural Language Processing(NLP). In this system the audio is given as input to our system instead of the text input. The system is designed in such a way that the audio input is converted to text for the further processing.

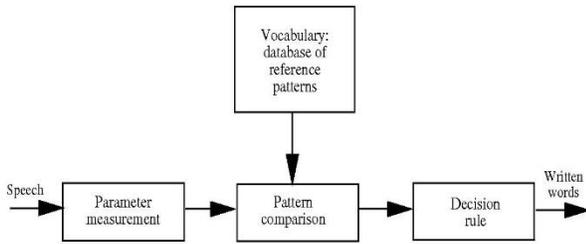


Fig No .1 Basic Working of Text To Speech

In the processing part the input is filtered that only specific keywords will act as the input and all the other words will be discarded this done by filtering the words. A database of the specific keywords is prepared and stored. The audio input will act as the query and find the matching content from the database and if no content is found it will show the message content not available or result not found and it will ask the user to give the audio input again.

VII. SYSTEM ARCHITECTURE

The system architecture we propose to use is as shown in the figure below. The various components and the modules used to develop the speech based food ordering system are classified below :

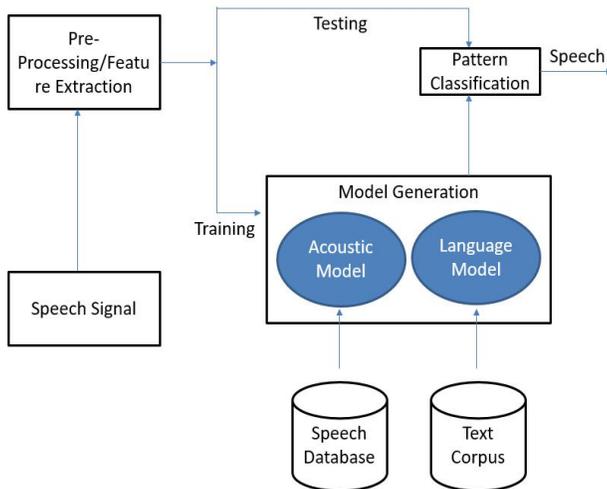


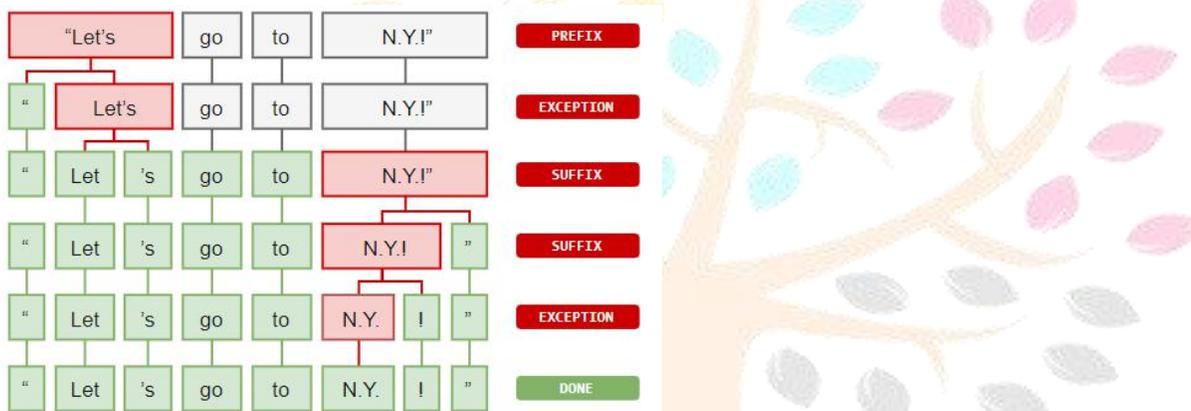
Fig No .2 System Architecture

The various components used in the system are pre-processing/feature extraction then the signal goes to testing and training components in the training part the various different types are stored then which checked in the pattern matching with the input speech and the output is given.

VI. DETAILED WORKING OF SPACY

1. **Tokenization:** Segmenting text into words, punctuations marks etc.
2. **Part-of-speech(POS) Tagging:** Assigning word types to tokens, like verb or noun.

3. **Dependency Parsing:** Assigning syntactic dependency labels, describing the relations between individual tokens, like subject or object.
4. **Sentence Boundary Detection (SBD):** Finding and segmenting individual sentences.
5. **Named Entity Recognition (NER):** Labelling named “real-world” objects, like persons, companies or locations.
6. **Similarity:** Comparing words, text spans and documents and how similar they are to each other.
7. **Text Classification:** Assigning categories or labels to a whole document, or parts of a document.
8. **Rule-based Matching:** Finding sequences of tokens based on their texts and linguistic annotations, similar to regular expressions.
9. **Training:** Updating and improving a statistical model’s predictions.
10. **Serialization:** Saving objects to files or byte strings.



VII. CONCLUSION

The conclusion is obtained that a speech based food ordering system can be developed. This will take speech as the input in the system will convert it into text. This system will be used by the various people of different age groups and also it can be used by the less tech savvy people who have little idea of the technical devices.

REFERENCES

- [1] "Bing Speech API," [online] <https://azure.microsoft.com/enus/services/cognitive-services/speech>, Jan. 2018
- [2] J. Adorf, "Web Speech API," Technical Report, KTH Royal Institute of Technology, Jan. 2018. https://www.researchgate.net/publication/327517552_Sinhala_Speech_Recognition_for_Interactive_Voice_Response_Systems_Accessed_Through_Mobile_Phones
- [3] "Nuance Recognizer," [online] <https://www.nuance.com/omnichannel-customer-engagement/voice-and-ivr/automatic-speechrecognition/nuance-recognizer>, Jan. 2017.
- [4] Dunn, Michael, Pro Microsoft Speech Server2007: Developing Speech Enabled Applications with. NET (Pro). Apress, 2007.