Voice enabled user interface

Demo:

https://keshavnagpal.github.io/nlp

July 8 2016

Contents

Objective

To make a proof of concept which can take voice input and process user's query provided in natural language to show desired outcome based on that query

Requirements

1. Microphone and permission to use it to take speech input.
2. A JavaScript enabled web browser supporting html5 web speech API.

Features and how to use

1. Users can use their voice to provide inputs to the interface
2. The query can be provided in natural language
3. To input a query user needs to click on the microphone button and start speaking, when no voice is detected user will be shown the respected outcome.
4. In case of any action which is not supported, a voice message will be provided to the user
5. Sample queries are integrated in the web page to help the user interact with the interface
6. Cascading of queries works upto two levels
7. The processing of queries can also handle corrections if user speaks something not meant as a query then user can speak the corrected

query or can also modify the query and alter the input and only the latest inputs will be taken as valid

8. User can select from different dialects of English language

Example/use cases

Speech recognition and natural language processing can be used to change the de-facto way of user interactions. Functions can be triggered using speech input, text can be extracted from speech and processed to create and provide specific outcomes for the user.

Logic

Speech Recognition

Used HTML5 web speech API and JavaScript functions to convert speech to text then that text can be fetched in processing the query

Natural Language Processing

The text converted from the speech input is processed by JavaScript functions made specifically for the desired requirements. One function extracts the keywords by matching them with supported actions, objects and their states. Data is then fetched to the respected action function which then performs the action and filters the result according to the object and its desired state provided by the user.

Cascading filters

Whenever user passes a query starting with keywords like “which of them “, ”from these how many” etc. then the function uses the result of previous query as the sample space for the current one

Technologies and frameworks used

1. HTML5 web speech API and JavaScript functions to convert text to speech
2. Materialize.css framework with HTML and CSS to design the Interface
3. jQuery ready function to initialize the data

Ideas for future

1. Add and Delete functions can also be integrated to the application to add and delete capabilities on voice commands.
2. More filters can be integrated.
3. We can expand the grammar supported.
4. More assisting voice messages can be added to help the user understand the process more efficiently.

Things to consider for further developments

1. To add more functions like add and delete, the action list must be updated with appropriate keywords and separate functions for each action needs to be created
2. The voice message spoken to the user when an unsupported action is passed in the query is using an online API not free for commercial use, it need to be removed or some alternate must be added before any commercial use