INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

Voice Based Search Engine for Visually Impaired People

Aravinthan T[1], Sabari A B[2], Someash Keerthi N S[3], William Roose D[4], S. Gopi[5]

UG Student, Dept. of I.T., Panimalar Engineering College., Chennai, Tamilnadu, India[1,2,3,4]

Assistant Professor, Dept. of I.T., Panimalar Engineering College., Chennai, Tamilnadu, India[5]

ABSTRACT: World Wide Web (WWW) is rapidly emerging as the universal information source for our society. The design of information on the web is visually oriented. The reliance on visual presentation places high cognitive demands on a user to operate such a system. The focus of this project is to develop a prototype which supports web browsing using a speech-based interface, e.g. a phone, and to measure its effectiveness. The command input and the delivery of web contents are entirely in voice. Audio icons are built into the prototype so that users can have better understanding of the original structure/intent of a web page. Navigation and control commands are available to enhance the web browsing experience. Voice browsers allow people to access the Web using speech synthesis, pre-recorded audio, and speech recognition. This can be supplemented by keypads and small displays. Voice may also be offered as an adjunct to conventional desktop browsers with high resolution graphical displays, providing an accessible alternative to using the keyboard or screen, for instance in automobiles where hands/eyes free operation is essential.

KEYWORDS: Data Mining; Voice Recognition; Speech to Text Conversion; Database Connectivity; Speech synthesis;

I. INTRODUCTION

Data Mining is the process of posing queries to large amounts of data sources and extracting patterns and trends using statistical and machine learning techniques. It integrates various technologies including database management, statistics and machine learning. Data mining has applications in numerous disciplines including medical, financial, defence and intelligence. Data mining tasks include classification, clustering, making associations and anomaly detection.

Database, data warehouse, World Wide Web (WWW), text files and other documents are the actual sources of data. You need large volumes of historical data for data mining to be successful. Organizations usually store data in databases or data warehouses. Data warehouses may contain one or more databases, text files, spreadsheets or other kinds of information repositories. Sometimes, data may reside even in plain text files or spreadsheets. World Wide Web or the Internet is another big source of data.

Regression technique can be adapted for predication. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining independent variables are attributes already known and response variables are what we want to predict. Unfortunately, many real-world problems are not simply prediction. For instance, sales volumes, stock prices, and product failure rates are all very difficult to predict because they may depend on complex interactions of multiple predictor variables. Therefore, more complex techniques (e.g., logistic regression, decision trees, or neural nets) may be necessary to forecast future values. The same model types can often be used for both regression and classification.

II. RELATED WORK

Association and correlation is usually to find frequent item set findings among large data sets. This type of finding helps businesses to make certain decisions, such as catalogue design, cross marketing and customer shopping behavior analysis. Association Rule algorithms need to be able to generate rules with confidence values less than one. However the number of possible Association Rules for a given dataset is generally very large and a high proportion of the rules are usually of little (if any) value. Types of association rule

- Multilevel association rule

- Multidimensional association rule

Neural network is a set of connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights so as to be able to predict the correct class labels of the input tuples. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. These are well suited for continuous valued inputs and outputs. For example handwritten character reorganization, for training a computer to pronounce English text and many real world business problems and have already been successfully applied in many industries. Neural networks are best at identifying patterns or trends in data and well suited for prediction or forecasting needs. Types of neural networks

- Back Propagation

III. PROPOSED ALGORITHM

Voice search, also called voice-enabled search, allows the user to use a voice command to search the Internet, or a portable device. Currently, voice search is commonly used in (in a narrow sense) "directory assistance", or local search. In a broader definition, voice search include open-domain keyword query on any information on the Internet. Voice search is often interactive, involving several rounds of interaction that allows a system to ask for clarification. Voice search is a type of *dialog system*. Voice search is a speech recognition technology that allows users to search by saying terms aloud rather than typing them into a search field. The proliferation of smart phones and other small, Web-enabled mobile devices has spurred interest in voice search. Applications of voice search include:

- Making search engine queries.

- Clarifying specifics of the request.

- Requesting specific information, such as a stock quote or sports score.

- Launching programs and selecting options.

The free voice search service, however, uses another approach. It might seem obvious, but people search differently using voice than when they type in a query. Speech recognition and generation technologies offer a potential solution to these problems by augmenting the capabilities of a web browser. The user can speak with the computer and the computer will respond to the user in the form of voice. The computer will assist the user in reading the documents as well.

IV. PSEUDO CODE

Step 1: User has to generate voice through microphone.
Step 2: Recognised voice should be noise eliminated using HMM
Step 3: Below Formula given the HMM model to convert speech into text

$$P(q1,\ldots,qn|x1,\ldots,xn) = \frac{P(x1,\ldots,xn|q1,\ldots,qn)P(q1,\ldots,qn)}{P(x1,\ldots,xn)}$$

$$P(q1, \ldots, qn) = \prod_{i=1}^{n} P(qi\,|qi-1).$$

Step 4: Given formula convert the speech into text.
Step 5: It connects with the server database using Google search engine.
Step 6: Displaying the relevant links on browser
Step 7: Again it coverts speech into text and go to step 2.
Step 8: End.

V. RESULTS

Voice search, also called voice-enabled search, allows the user to use a voice command to search the Internet, or a portable device. Currently, voice search is commonly used in (in a narrow sense) "directory assistance", or local search. In a broader definition, voice search include open-domain keyword query on any information on the Internet. Voice search is often interactive, involving several rounds of interaction that allows a system to ask for clarification. Voice search is a type of *dialog system*. Voice search is a speech recognition technology that allows users to search by saying terms aloud rather

than typing them into a search field. The proliferation of smart phones and other small, Web-enabled mobile devices has spurred interest in voice search. Applications of voice search include:

- Making search engine queries.
- Clarifying specifics of the request.
- Requesting specific information, such as a stock quote or sports score.
- Launching programs and selecting options.

The free voice search service, however, uses another approach. It might seem obvious, but people search differently using voice than when they type in a query. Speech recognition and generation technologies offer a potential solution to these problems by augmenting the capabilities of a web browser. The user can speak with the computer and the computer will respond to the user in the form of voice. The computer will assist the user in reading the documents as well.
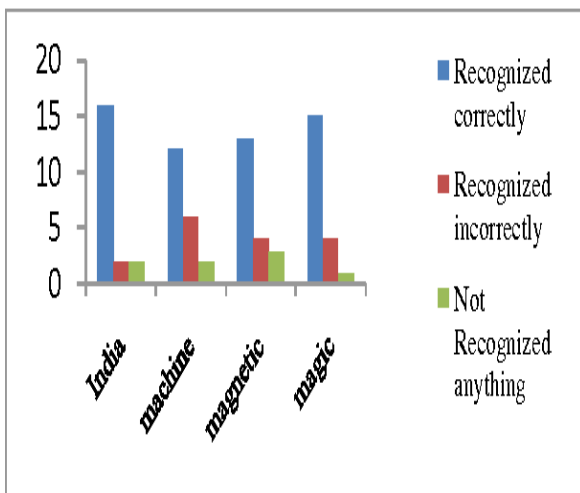

C-4	10

TABLE 3
ACCURACY IN PERCENTAGE

Combination No.	Accuracy (%)
C-1	80
C-2	90
C-3	100
C-4	90

Fig.1. Default Words Taken

Fig. 2 Word Comparison Accuracy

TABLE 4
ACCURACY MEASURES

Combination No.	Sample words	No. of time recognized correctly	No. of times Recognized incorrectly	No. of times not Recognized
C-1	India	16	4	0
C-2	Machine	18	2	0
C-3	Magnetic	20	0	0
C-4	Magic	18	2	0

Fig. 3. Word Comparison Results

VI. CONCLUSION AND FUTURE WORK

In this project we proposed an efficient way of accessing the web browser is presented which is termed as voice browsing in which visually impaired people can access the browser using speech. As access to internet visually incurs limitations such as visually impaired persons cannot use keypads, touch screens etc. for giving inputs to computer. User can speech the word and converted into text automatically, now this browser reduces their effort by performing this conversion automatically. And the blind people can also use this browser to convert text documents in English characters. Thus combination of browsing with speech technology is an efficient way of accessing webs. This methodology can be further improvised for a browser that allows visually impaired learners to interact more efficiently with the browser by converting their English characters to speech i.e. listening of characters, which can be easily understood by them. In addition, all the text content present over the web for various links can be made accessible by using speech technology. This technology can also be implemented in browser. More work can be done to increase the accuracy, pronunciation and precision of speech technology. The proposed method has used only English language. In future, we can extend the framework to implement various languages and also various web access contents with improved accuracy

REFERENCES

[1] Furui, Sadaoki. "Recent progress in corpus-based spontaneous speech recognition." IEICE Transactions on Information and Systems 88.3 (2005): 366-375.

[2] Hinton, Geoffrey, et al. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups." IEEE Signal processing magazine 29.6 (2012): 82-97.

[3] Real, Esteban, et al. "Large-scale evolution of image classifiers." arXiv preprint arXiv:1703.01041 (2017).

[4] Snoek, Jasper, Hugo Larochelle, and Ryan P. Adams. "Practical bayesian optimization of machine learning algorithms." Advances in neural information processing systems. 2012.

[5] Shahriari, Bobak, et al. "Taking the human out of the loop: A review of bayesian optimization." Proceedings of the IEEE104.1 (2016): 148-175.

[6] Snyder, David, Daniel Garcia-Romero, and Daniel Povey. "Time delay deep neural network-based universal background models for speaker recognition." Automatic Speech Recognition and Understanding (ASRU), 2015 IEEE Workshop on. IEEE, 2015.

[7] Dahl, George E., Tara N. Sainath, and Geoffrey E. Hinton. "Improving deep neural networks for LVCSR using rectified linear units and dropout." Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on.IEEE, 2013.

[8] Watanabe, Shinji, and Jonathan Le Roux. "Black box optimization for automatic speech recognition." ICASSP. 2014.

[9] Shinozaki, Takahiro, and Shinji Watanabe. "Structure discovery of deep neural network based on evolutionary algorithms." Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on. IEEE, 2015

[10] Chandrashekaran, Akshay, and Ian Lane. "Automated optimization of decoder hyper-parameters for online lvcsr." Spoken Language Technology Workshop (SLT), 2016 IEEE.IEEE, 2016.