e-ISSN 2320 –7876 www.ijfans.org
Vol.9, Iss.2, Feb 2020
 UGC CARE Listed (Group -I) Journal

A NOVEL REAL TIME VOICE BASED APPROACH FOR MULTILINGUAL WEB DATA EXTRACTION WITH RASPBERRY PI

[1]ANJANEYA VARAPRASAD SIRIGINEEDI, [2]SATYA VENI YAMANA, [3]PALLI SRINIVAS, [4]YALLA NAGA VENKATA SAILAJA

[1,2]Assistant Professor, Department of E.C.E, BVC College of Engineering, Rajahmundry, A.P, India.

[3]Associate Professor, Department of E.C.E, BVC College Of Engineering, Rajahmundry, A.P, India.

[4]B. Tech Scholar, Department of E.C.E, BVC College of Engineering, Rajahmundry, A.P, India.

ABSTRACT: The Web has witnessed an enormous growth in the amount of semantic information published in recent years. This growth has been stimulated to a large extent by the emergence of Linked Data. Although this brings us a big step closer to the vision of a Semantic Web, it also raises new issues such as the need for dealing with information expressed in different natural languages. As a consequence, the need to extract useful information from different languages increases, highlighting the importance of research into Open Information Extraction (OIE) techniques. Different OIE methods have dealt with features from a unique language; however, few approaches tackle multilingual aspects. In those approaches, multilingualism is restricted to processing text in different languages, rather than exploring cross-linguistic resources, which results in low precision due to the use of general rules. For multilingual users, a further obstacle to natural interaction is the monolingual character of many speech recognition systems, in which users are constrained to a single preset language. In order to solve these issue, a novel real time voice based approach for Multilingual web data extraction with raspberry pi is presented in this work. The main aim of this work is to assist the users (student) to get accurate data. If the user wants to search some data in different languages the user need not to search multiple times in different languages. That means if user gives one language as input it displays the result in multiple languages as output.

KEYWORDS: Multilingual, Semantic Web data, Voice, Speech recognition and Raspberry Pi

I. INTRODUCTION

Today we live in the information age, technology specifically like the web and the internet, in particular, has changed the way we work and communicate.

There is an enormous amount of information available for everyone. However, the amount of Information available is no use if there is not a suitable technique to process the information and extract knowledge from information. Text mining is one of the technologies that apply for those purposes which designed specially to deal with unstructured data. To make easier for a user to access the extracted information easily and effectively without and extra effect speech is applied. Speech now days are used for accessing any information. To do that user just interacts with the system by voice command and system recognizes the voice and invokes the necessary module for the output [3].

Textual data are the main form of data published in the Web, and the number of published documents increases daily. As much as the Web is a valuable source of information and knowledge, the sheer amount of available pages renders it impossible for a person to explore all of the available information on any subject. That massive amounts of semantic data are becoming available on the Web, the question emerges how end users should access and interact with this wealth of data. As language is the most important means by which humans communicate, it is reasonable to assume that users would find a language-mediated way of accessing the Web of Data intuitive, appealing and effortless. In fact, the traditional Web is language-specific and information can only be accessed across languages if web

sites are translated into the corresponding languages.

The language diversity and lack of technological support for spoken languages in India make universal access to information and services an ongoing challenge. Automatic speech recognition (ASR) has become increasingly relevant to date, tracking the explosive growth of mobile devices. The use of voice as a natural and convenient method of human-device interaction is especially applicable to hands-free scenarios (e.g., while driving) and interaction with small form-factor devices (e.g., wearables). The quality of the user experience in these scenarios is primarily affected by the transcription accuracy and real-time responsiveness of the ASR system [6].

There have been several successful attempts in the development of systems that can analyze, classify and recognize speech signals. Both hardware and software that have been developed for such tasks have been applied in various fields such as health care, government sectors and agriculture [9]. Speaker recognition is the capability of a software or hardware to receive speech signal, identify the speaker present in the speech signal and recognize the speaker afterwards. Speaker recognition executes a task similar to what the human brain undertakes.

Information Extraction is the process of accessing required information from unstructured data in short amount of time. Information extraction (IE) process extracts useful structured information from the unstructured data in the form of entities, relations, objects, events and many other types. The extracted information from unstructured data is used to prepare data for analysis. Therefore, the efficient and accurate transformation of unstructured data in the IE process improves the data analysis. Numerous techniques have been introduced for different data types i.e. text, image, audio, and video. Due to the huge volume and complexity of unstructured data, it became a tedious task to extract useful information from different types of data. Several architectures have been considered to achieve multilingual data extraction.

However they have complexities due to large volumes of unstructured data. Hence in order to solve these issues, a novel real time voice based approach for multilingual web data extraction with Raspberry Pi is presented. The rest of the work is organized as follows: The section II describes literature survey. The section III demonstrates a novel real time voice based approach for multilingual web data extraction with Raspberry Pi. The section IV presents the result analysis of presented approach. The conclusion is provided in section V.

II. LITERATURE SURVEY

Ziqi Zhang, David Robinson, and Jonathan Tepper et. al., [1] presents Detecting Hate Speech on Twitter Using a Convolution-GRU Based Deep Neural Network. This work introduces a new method based on a deep neural network combining convolutional and gated recurrent networks. Authors conducted an extensive evaluation of the method against several baselines and state of the art on the largest collection of publicly available Twitter datasets to date, and show that compared to previously reported results on these datasets, described method is able to capture both word sequence and order information in short texts, and it sets new benchmark by outperforming on 6 out of 7 datasets by between 1 and 13% in F1.

Pablo Gamallo, Marcos Garciay, C´esar Pi˜neiro, Rodrigo Mart´ınez-Casta˜no and Juan C. Pichel et. al., [2] describes LinguaKit: a Big Data-based multilingual tool for linguistic analysis and information extraction. This work presents LinguaKit, a multilingual suite of tools for analysis, extraction, annotation and linguistic correction, as well as its integration into a

Big Data infrastructure. LinguaKit allows the user to perform different tasks such as PoS-tagging, syntactic parsing, coreference resolution (among others), including applications for relation extraction, sentiment analysis, summarization, extraction of multiword expressions, or entity linking to DBpedia. Most modules work in four languages: Portuguese, Spanish, English, and Galician. The system is programmed in Perl and is freely available under a GPLv3 license.

Kanwal Yousaf, Zahid Mehmood , Tanzila Saba, Amjad Rehman, Muhammad Rashid, Muhammad Altaf, and Zhang Shuguang et. al., [3] describes A Novel Technique for Speech Recognition and Visualization Based Mobile Application to Support Two-Way Communication between Deaf-Mute and Normal Peoples" In this work mel frequency cepstral coefficients (MFCC) based features are extracted for each training and testing sample of Deaf-mute speech. The hidden Markov model toolkit (HTK) is used for the process of speech recognition. The application is also integrated with a 3D avatar for providing visualization support. The quantitative and qualitative analysis of results also revealed that face-to-face socialization of Deaf-mute is improved by the intervention of mobile technology. The participants also suggested that the described mobile application can act as a voice for them and they can socialize with friends and family by using this app.

Sagar Patil, Mayuri Phonde, Siddharth Prajapati, Saranga Rane, Anita L ahane et. al., [5] presents Multilingual Speech and Text Recognition and Translation using Image. This application recognizes speech (human matter) in one language to another user defined language to communicate in expressive manner. It includes 4 modules voice recognition, translation and speech synthesis and image translation and gives audio of the translated language. Also the application accepts text written and converts it into the language needed Application is able to recognize the text present in the image which stored in system or captured using camera and translate the text into the language needed and display the translation result back on to the screen of system.

Hanna Suominen, Liyuan Zhou, Leif Hanlen, Gabriela Ferraro et. al., [7] presents Benchmarking Clinical Speech Recognition and Information Extraction. The objective of the study was to provide a recorded spoken handover, annotated verbatim transcriptions, and evaluations to support research in spoken and written natural language processing for filling out a clinical handover form. This dataset is based on synthetic patient profiles, thereby avoiding ethical and legal restrictions, while maintaining efficacy for research in speech-to-text conversion and information extraction, based on realistic clinical scenarios. They also introduce a Web app to demonstrate the system design and workflow.

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, S¨oren Auer, Christian Bizer et. al., [10] presents DBpedia A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. The DBpedia community project extracts structured, multilingual knowledge from Wikipedia and makes it freely available on the Web using Semantic Web and Linked Data technologies. The project extracts knowledge from 111 different language editions of Wikipedia. In this system report, they give an overview of the DBpedia community project, including its architecture, technical implementation, maintenance, internationalization, usage statistics and applications.

III. NOVEL REAL TIME VOICE BASED APPRAOCH FOR

MULTILINGUAL DATA EXTRACTION

In this section, a novel real time voice based approach for multilingual web data extraction with Raspberry Pi is presented. The block diagram of presented approach is shown in Fig. 1.

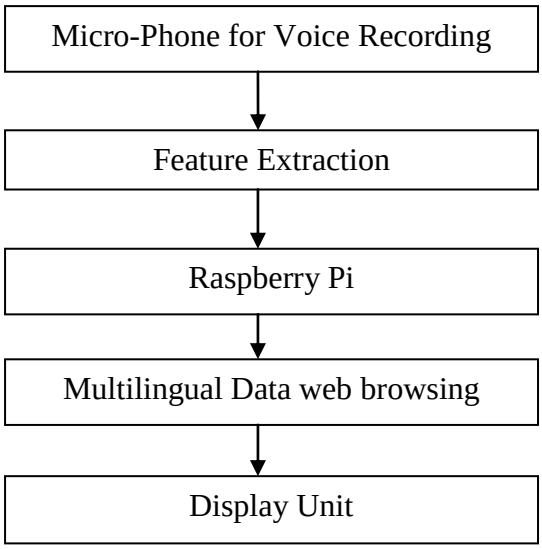

Fig. 1: Block Diagram of Novel real time voice based approach for multilingual web data extraction with Raspberry Pi

Microphone is an input instrument for converting sound waves into electrical energy variations which may then be amplified, transmitted or recorded Initiation of speech signal is produced from the air-flow of lungs that creates vibration to the vocal folds (called vocal cords). The produced time-varying waveform which are the results of actions from respiratory and articulatory organs is called speech. The voice organ or our phonetic system belongs to three different systems- these are breathing apparatus, vocal folds and vocal tract. Vocal folds is the most important organ to produce sound because based on the vibration of vocal folds or vocal cords, voiced and unvoiced sounds are created. Basically, all vowels are included as voiced sounds, there also some consonants which are also voiced.

The speech signal has to be processed to remove noise before the extraction of the important attributes in the speech and identification. The purpose of feature extraction is to illustrate a speech signal by a predetermined number of components of the signal. This is because all the information in the acoustic signal is too cumbersome to deal with, and some of the information is irrelevant in the identification task.

A keyboard is one of the primary input devices used to enter text. Mouse is also an input device that allows you to control the coordinates and movement of the onscreen cursor/pointer. Ethernet is being used to provide internet connection to the voice command. For online query processing, constant connection of internet is needed. The Raspberry Pi needs a constant 5v,1.2 Ma power supply.

In this project Raspberry-Pi3 is a main tool that all other components are get connected to it. Raspberry-Pi3 which is a microprocessor(acts as a small computer) will process all the data in it and gives the appropriate result as output. Actually we dump the code into Rasberry-Pi3 which we want to executes to get the required output. one can run the program by using Rasberry-pi3 and get the output. As compared to other Micro processor Raspberry-Pi is very small in size and even it is smaller than a mobile phone and it is just like a debit card in size. The main parts used in Raspberry pi are: HDMI port, USB ports, Micro SD card, Memory and Power which are discussed as follows:

HDMI (High-Definition Multimedia Interface): HDMI is a compact audio and video interface for transmitting digital data in single cable It carries the audio and video streams together , eliminating cable clutter and the information can be transmitted digitally using this HDMI cable. It contains three physically separate communication channels ,which are the DDC , TMDS , and the optional CEC.

Universal Serial Bus(USB): USB is type of computer port which can be used to

connect equipment to a computer. It is used to designate a kind of stand interface for connecting peripherals to a computer. A USB is a serial bus with a data transfer rate of 12mbps for connecting a micro computer to peripherals such as keyboards, mice, printers, and digital cameras through a single , general purpose port.

Memory: It uses a 1.2GHZ 64 Bit quad-core Arm Cortex-A53 CPU, has 1GB RAM , integrated 802.11N wireless LAN, & Bluetooth 4.1. This puts the pi3 roughly 50% faster than the pi2. Micro Sd Port (Micro Secrue Digital Slot): It is a small explanation slot located in mobile and portable devices. It facilitates the increase of available memory via the insertion of a MicroSD card. MicroSD slots are built for portable devices that require storage, such as mobile phones, tablets, and digital cameras. Many personal computers (PC) and laptops are built with SD card slots but with larger devices, a slot only accepts an SD card. Power the raspberry pi is via the micro port USB on the side of the unit. The recommended input voltage is 5volts and recommended current is 2A. The latest raspberry pi recommends 5 volts@ 2A as a minimum for stability, but some are more flexible.

Monitor is a display unit of computer. It is main output device that displays the processed data as text and images. Speakers are transducers that convert electromagnetic waves into sound waves.

Initially the input is given either by using microphone (speech as input) or by using keyboard (text as input). If the input is given by using microphone through speech then first it will recognized by the raspberry-Pi3 and then it will converted into text. After that processor (Raspberry-Pi) will go the web browsing and search for the required output that the user want.

IV. RESULT ANALYSIS

In this section, a novel real time voice based approach for multilingual web data extraction with Raspberry Pi is implemented. The main modules in this approach are microphone, Raspberry Pi, Display unit, etc. the input is given either by voice or through keyboard. The raspberry pi recognizes the voice and converts it into text. Next it browses the multilingual data and provides the data in multiple languages.

The Fig. 2 shows the output of presented approach.

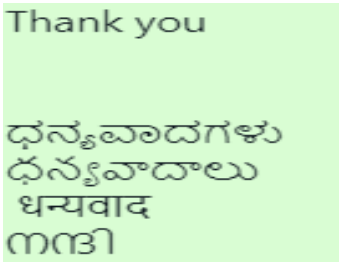

Fig. 2: Results Screen

If the user enters Thank you through voice or keyboard then the output is displayed in multiple languages. In this manner this approach displays multilingual data based on Voice signal. This approach provides Real time spoken translations and is used to develop the communication skills of the user. As compared with GOOGLE or other web browser this project will give the result of the searched query in different languages at a time without browsing the query in multiple times as others can't.

V. CONCLUSION

A novel real time voice based approach for multilingual web data extraction with Raspberry Pi is implemented in this work. The main objective of this work is to assist the users (student) to get accurate data. The main modules in this approach are microphone, keyboard, Raspberry Pi, Raspberry pi modules, Display unit. Microphone is used to record the user voice data for further process. The Raspberry pi process the voice signal and

decodes the voice data and displays it in multiple languages. The output is displayed on the computer monitor, in addition output is delivered through the speakers. This approach Process the voice/text as an input. Next Decodes the input command and Gives the appropriate relevant output in different languages. With this approach, the user doesn't need to search multiple times in different languages when he wants to search certain data in different languages. This approach has provided Real time spoken translations and is used for developing the communication skills of the users.

VI. REFERENCES

[1] Ziqi Zhang, David Robinson, and Jonathan Tepper, "Detecting Hate Speech on Twitter Using a Convolution-GRU Based Deep Neural Network", Springer International Publishing AG, part of Springer Nature 2018, ESWC 2018, LNCS 10843, pp. 745–760, 2018, doi:10.1007/978-3-319-93417-4_48

[2] Pablo Gamallo, Marcos Garciay, C´esar Pi˜neiro, Rodrigo Mart´ınez-Casta˜no and Juan C. Pichel, "LinguaKit: a Big Data-based multilingual tool for linguistic analysis and information extraction", 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS),
DOI: 10.1109/SNAMS.2018.8554689

[3] Kanwal Yousaf , Zahid Mehmood , Tanzila Saba, Amjad Rehman, Muhammad Rashid, Muhammad Altaf, and Zhang Shuguang, "A Novel Technique for Speech Recognition and Visualization Based Mobile Application to Support Two-Way Communication between Deaf-Mute and Normal Peoples", Hindawi Wireless Communications and Mobile Computing Volume 2018, Article ID 1013234, 12 pages, doi:10.1155/2018/1013234

[4] Amit Verma, Iqbaldeep Kaur, Amit Kumar, "Information Extraction with Speech Recognition", International Journal of Computer Science and Information Security (IJCSIS), Vol. 15, No. 4, April 2017, ISSN 1947-5500

[5] Sagar Patil, Mayuri Phonde, Siddharth Prajapati, Saranga Rane, Anita L ahane, "Multilingual Speech and Text Recognition and Translation using Image", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 5 Issue 04, April-2016

[6] Javier Gonzalez-Dominguez, David Eustis, Ignacio Lopez-Moreno, Andrew, Françoise Beaufays, and Pedro J. Moreno, "A Real-Time End-to-End Multilingual Speech Recognition Architecture", IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 9, NO. 4, JUNE 2015

[7] Hanna Suominen, Liyuan Zhou, Leif Hanlen, Gabriela Ferraro, "Benchmarking Clinical Speech Recognition and Information Extraction: New Data, Methods, and Evaluations", JMIR MEDICAL INFORMATICS, 2015;3(2):e19) doi: 10.2196/medinform.4321

[8] R. Sousa, A. Ferreira, and P. Alku, "The harmonic and noise information of the glottal pulses in speech", Biomedical
Signal Processing and Control, Elsevier, vol. 10, pp. 137–143, 2014.

[9] Ezeiza A, Lopez-de-Ipin˜a K, Herna´ndez C, Barroso N, "Enhancing the feature extraction process for automatic speech recognition with fractal dimensions", Cogn Comput 5(4):545–550, 2013.

[10] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, S¨oren Auer, Christian Bizer, "DBpedia A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia", Semantic Web 1 (2012) 1–5 1
IOS Press, 1570-0844/12, 2012

[11] S. B. Jaramillo, E. B. Bolanos, T. V. Canas, J. R. O. Arroyave, J. D. A. Londono, and J. F. V. Bonilla, "Automatic emotion detection in speech using mel frequency cesptral coefficients", In Proc. of Symposium of Image, Signal Processing, and Artificial Vision, IEEE, Antioquia, Colombia, 2012

[12] Lopez de Ipin˜a K, Torres I, On˜ederra L, Varona A, Ezeiza N, "First selection of lexical units for continuous speech recognition of Basque", In: Proceedings of ICSLP, vol 2, pp 531–535. Beijing