Volume-6, Issue-4 (Nov-16)
ISSN (O) :- 2349-3585

Review on Punjabi Question Answer

Anjani Garg[1], Bharti Grover[2]
[1,2] Department of Computer Science Engineering
[1,2] GVIET, Ramnagar,Banur, Punjab, India.

Abstract- In recent years, Question Answering (QA) system has been researched extensively and automatic question answering has become an interesting research field and resulted in a visible improvement in its performance. Especially, during the last decade, a number of automatic QA systems have emerged. There are various approaches that can be used to generate the questions from a given text. In this paper we are presenting the review on question generation from documents written in Punjabi language.

Keywords – Punjabi, Question, Answer, Language.

INTRODUCTION

As the requirement of information is essential part of our life[1]. There are numerous sources of information, but the major one is database. Database helps us to store, access and retrieve information. No organization or industry is possible without the use of database. Each and every computer based application need to access information from database that requires knowledge of formal query language like SQL. But it is not possible for everyone to learn or write SQL queries. To overturn this problem many researchers have brought out to use Natural Language (NL) i.e. Punjabi, Arabic, English, Bengali etc. in place of formal query language which can be a perfect interface between an application of computer and nontechnical user. This idea of using NL has induced the development of new sort of processing method in database systems. From the last decade internet [2] users are increasing in more titre. In QA different queries are provided by the user in aim of getting accurate answers in Question Answering Systems. Question Answering provides perfect solution to retrieve valid and accurate answers to user question asked in natural language instead of query. Various Scientists are Work on QA from last many with the English, Chinese, Japanese, Korean, etc. languages. But from some last year's scientists move their research into regional languages like Tamil, Punjabi, Hindi, Malayalam, etc. In today's context[13], there are three components of Question answering system (QAS) are question classification, information retrieval, and answer extraction. These components play a vital role in QAS. Question classification play chief role in QA system to categorize the question based upon its type. Information retrieval method is to identify success by extracting out applicable answer post by their intelligent question answering system. In comparing with classical information retrieval, where whole documents are considered similar to the information demand, in question answering, specific pieces of information are returned as an answer. The user of a question answering system is interested in a concise, comprehensible, correct and most probable answer, which may refer to a word, sentence, paragraph, image, audio fragment, or an entire document. Question answering system is an important application of the text mining. These systems extract more relevant answers to a particular question posed to them from the large collection of text corpus. On the other hand, various relevant answers can also be extracted classifying the questions into various categories and hence relevancy can be detected easily from the questions and the answers. The Internet today has to face the difficulty of dealing with multi languish. All the work in Question answering system is done for various other languages but as per knowledge, limited work is done for Punjabi Language.

LITERATURE REVIEW

Saravjeet Kaur et al. (2012) [3] proposed a method of querying with the databases by means of a natural language interface. This is hot issue in the area of database management is to provide a high level interface for nontechnical users. Normal users are not aware with the formal language like SQL. Then the problem is how they interact with the database system. A normal user may find him/her self handicapped to deal with the database system. The paper presents an interface

module that converts user's query given in natural language into a corresponding SQL command. Asking questions to databases in natural language like English is a very convenient and easy method of data access from database system, especially for normal users who do not understand complicated database query languages such as SQL. This paper proposed the architecture for translating English Query into SQL. Shriya Sahu1 et al. (2012) [4] presents an approach to extract answers from Hindi text for a given question. It is based on understanding the meaning of the given question and expressing them in query logic language. The Hindi text is analyzed to understand the semantic of each sentence and relevant answer is extracted for the given question. The answers are extracted for the questions of type when, where, how many and what time.The developed Question-Answering system in Hindi uses Hindi Shallow Parser which is developed by IIIT Hyderabad. The shallow parser gives the analysis of a sentence in terms of morphological analysis, POS tagging, Chunking, etc. Apart from the final output, intermediate output of individual modules is also available. All outputs are in Shakti Standard Format (SSF). Wiqas Ghai et al., (2012) [5] discussed that Punjabi, Hindi, Marathi, Gujarati, Sindhi, Bengali, Nepali, Sinhala, Oriya, Assamese, Urdu are prominent members of the family of Indo-Aryan languages. These languages are mainly spoken in India, Pakistan, Bangladesh, Nepal, Sri Lanka and Maldive Islands. All these languages contain huge diversity of phonetic content. In the last two decades, few researchers have worked for the development of Automatic Speech Recognition Systems for most of these languages in such a way that development of this technology can reach at par with the research work which has been done and is being done for the different languages in the rest of the world. Punjabi is the 10th most widely spoken language in the world for which no considerable work has been done in this area of automatic speech recognition. Being a member of Indo-Aryan languages family and a language rich in literature, Punjabi language deserves attention in this highly growing field of Automatic speech recognition. In this paper, the efforts made by various researchers to develop automatic speech recognition systems for most of the Indo-Aryan languages, have been analysed and then their applicability to Punjabi language has been discussed so that a concrete work can be initiated for Punjabi language. Simarjeet Kaur et al., (2013) [6] discussed that the accurate and timely advice for the effective insect pest management is an important component in controlling the pest on vegetables crops. This web based insect pest management system aims to transfer the pest management practices in different vegetable crops recommended by Punjab Agricultural University to the farmers for their guidance to take quick and timely actions for pest management in their fields. This system has been developed using PHP, HTML, CSS, JavaScript and Ajax and database has been designed using MySQL. The proposed system is advantageous as it is easy to use, effective and efficient in controlling the insect pests by providing accurate and timely information at affordable cost. D.Ramesh et al. (2013) [7] discussed that the abundant information available on internet generates the need to store data in an organized manner so that searching, retrieving and maintaining of data becomes easier. A database is a technology that stores the data in an logical and organized manner. To efficiently operate these databases, knowledge of structures query language (SQL) becomes essential. But the usage of SQL restricts the access to databases from the users who don't have the knowledge of them. A need for interface comes into the picture to enable the access of these databases even to non-expert users. This paper describes the design to develop Telugu language interface to databases. Bank database is used as a case study to develop telugu language interface. The performance of the system has shown to be satisfactory. Ashish Kumar et al. (2013) [8] discussed that in the world of computing, information plays an important role in our lives. One of the major sources of information is database. Database and Database technology are having major impact on the growing use of computers. Almost all IT applications are storing and retrieving the information or data from the database. Database Management Systems (DBMS) have been widely used for storing and retrieving data. However, databases are often hard to use since their interface is quite rigid in co-operating with users. For storing and retrieving the information from database requires the knowledge of database language like SQL. Structured Query Language (SQL) is an

ISSN:-2349-3585 |www.ijrdt.org

ANSI standard for accessing and manipulating the information stored in database. However, everyone may not be able to write the SQL query as they may not be aware of the syntax and structure of SQL and database respectively. In India, the natural language of people is mainly Hindi. Also large number of e-governance applications use database. So, to use such database applications with ease, people who are more comfortable with Hindi language, requires these applications to accepts a simple sentence in Hindi, and process it to generate a SQL query. The SQL query is further executed on the database to produce the results. Therefore, any interface in Hindi language will be an asset to these people. This paper discusses the architecture of mapping the Hindi language query entered by the user into SQL query. Jaspreet Kaur et al. (2013) [9] discussed that as there is increasingly high advancement in technology, so Question Answering is becoming major area of research for the researchers. Different queries are provided by the user in aim of getting accurate answers in Question Answering Systems. Question Answering provides perfect solution to retrieve valid and accurate answers to user question asked in natural language instead of query. Hindi, Telugu, Bengali etc are popular languages that are spoken in India. Currently these languages are taken into consideration by the researchers and a lot of work is being done in these and other Indian languages. In this paper we compare Question Answering Systems performance for different Indian languages. We discuss the best features of Question Answering systems built in different Indian languages and compare their performances. Manu Bansal et al. (2013) [10] discussed that the term data mining has been the oldest yet one of the interesting buzzwords. Many organizations often underutilize their already existing databases. There is a need to mine information and interesting patterns from these databases. The focus of the current research is to apply data mining on a library management system. Data mining is usually done on a data warehouse or a data mart. It incurs various cost factors like software, hardware, maintenance and experts. The objective here is to study how the real-time data stored in database can be turned informative without setting up a separate data warehouse. The main emphasize is on understanding the problem perspective, competing objectives and constraints and generating a model for information extraction from the real-time library database using ARM (Association Rule Mining) mining technique. As SQL (Structured Query Language) can also be used for mining data instead of using specialized data mining algorithm, the study also compares SQL based mining with ARM. The results shows that association rule mining performs better than SQL based mining as type of pattern to be extracted can be controlled much effectively in ARM as compared to SQL because of the parameters (support and count) used in the data mining algorithm. Algorithms are implemented using SQL and MATLAB (Matrix Laboratory) Tool - ARMADA. Ravinder Kumar et al. (2014) [11] discussed that we require information in our daily life. One of the major sources of information is database. Almost all applications need to retrieve information from database that requires knowledge of database languages like SQL. To write SQL query one need to have knowledge of formal query language. Therefore everybody is not able to write SQL queries. To override the complexity many research have turned out to use Natural Language (NL) i.e. English, French, Tamil, Arabic, Hindi, Punjabi etc. instead of SQL. The idea of using NL has prompted the development of new type of processing method called Natural Language Interface to Database systems (NLIDB). Lots of works have been already done in natural language processing but in Hindi and other Indian languages have scope of research and improvement. So for this intention we are trying to design a tool and database that helps the farmers to solve out the different queries requested in their native language (Hindi) related to agriculture. To design the tool that fulfills the objectives of our work we will use Java swings as front end and for storing the data we will use MySQL 5.0 as backend. Priyanka Arora et al. (2014) [1] discussed that database Management Systems have been used extensively for accessing, storing and retrieving data. However, database systems are not understandable to every user because they are hard to use. A plethora of e-governance applications railways, billings, agriculture, banks etc. use databases. Some users face difficulty in using these database systems because they do not have knowledge of the languages used in this system. So, they want a system that accepts a

ISSN:-2349-3585 |www.ijrdt.org

Hindi sentence as s query and after processing it, execute it and provides the output in the same language. Then the users have no need to learn any low-level languages those are hard to learn, use in databases such as SQL. Rohini B. Kokare et al. (2014) [12] discussed that natural language query builder interface retrieves the required data from database when query is given in natural language. To retrieve the correct data from database, the user should have sufficient technical knowledge of Structured Query Language (SQL) statements. Natural Language Query Builder Interface (NLQBI) will solve this problem. In natural language parsing, getting highly accurate syntactic analysis is a crucial step. Parsing of natural languages can be seen as the process of mapping an input string or a sentence to its syntactic representation. One of the parsing techniques is dependency parsing. Dependency parsing focuses on relations between words which resolve ambiguity. Most of the recent efficient algorithms for dependency parsing work by factoring the dependency trees. Graph based dependency parsing models are prevalent in dependency parsing because of their state-of-art accuracy and efficiency. This paper covers some recent developments in NLQBI systems and survey on dependency parsing techniques. Preeti Verma et al. [4] discussed that unlike most user-computer interfaces, a natural language interface allows users to communicate fluently with a computer system with very little preparation. Databases are often hard to use in cooperating with the users because of their rigid interface. A good NLIDB allows a user to enter commands and ask questions in native language and then after interpreting respond to the user in native language. For a large number of applications requiring interaction between humans and the computer systems, it would be convenient to provide the end-user friendly interface. Punjabi language interface to database would proof fruitful to native people of Punjab, as it provides ease to them to use various e-governance applications like Punjab Sewa, Suwidha, Online Public Utility Forms, Online Grievance Cell, Land Records Management System, legacy matters, e-District, agriculture, etc. Punjabi is the mother tongue of more than 110 million people all around the world. According to available information, Punjabi ranks 10th from top out of a total of 6,900 languages recognized internationally by the United Nations. This paper covers a brief overview of the Natural language interface to database, its different components, its advantages, disadvantages, approaches and techniques used. The paper ends with the work done on Punjabi language interface to database and future enhancements that can be done.

CONCLUSION

In this paper we present the review to generate questions automatically from a given Punjabi text. In recent years, Question Answering (QA) system has been researched extensively and automatic question answering has become an interesting research field and resulted in a visible improvement in its performance. Especially, during the last decade, a number of automatic QA systems have emerged. In automation of question answers in Punjabi the input is taken from the user in Punjabi language.

REFERENCE

[1]. Priyanka Arora ,Punet Goswami, "An Eficient Hindi Language Interface using Relational Databases", International Journal of Research in Computer and Communication Technology, Vol 3, Issue 6, June – 2014.

[2]. Sunil A. Khillare, Bharat A. Shelke and C. Namrata Mahender(2014), "Comparative Study on Question Answering Systems and Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 11.

[3]. Saravjeet Kaur, Rashmeet Singh Bali, "SQL Generation and Execution from Natural Language Processing", International Journal of Computing & Business Research, Proceedings of 'I-Society 2012' at GKU, Talwandi Sabo Bathinda (Punjab).

[4]. Preeti Verma, Suket Arora, Kamaljit Batra, "Punjabi Language Interface to Database: A Brief Review".

[5]. Wiqas Ghai, Navdeep Singh, "Analysis of Automatic Speech Recognition Systems for Indo-Aryan Languages: Punjabi A Case Study", International Journal of Soft Computing and Engineering (IJSCE),ISSN: 2231-2307, Volume-2, Issue-1, March 2012.

[6]. Simarjeet Kaur, B K Sawhney, Sandeep Kaur, "Design and Development of Software for Insect Pest Management of Vegetable Crops using Web Technology", International Journal of Emerging Science and Engineering (IJESE) ISSN: 2319–6378, Volume-1, Issue-11, September 2013.

[7]. D.Ramesh, Suresh Kumar Sanampudi, "Telugu Language Interface to Databases", International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, Issue 7, July 2013.

[8]. Ashish Kumar and Kunwar Singh Vaisla, "Hindi Language Interface to Database using Semantic Matching", Oriental Journal of Computer Science & Technology, June 2013, vol. 6, no. (2):Pgs.133-140.

[9]. Jaspreet Kaur & Vishal Gupta, "Comparative Analysis of Question Answering System in Indian Languages", 128X International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, July 2013.

[10]. Manu Bansal and Mandeep Kaur, "Analysis and Comparison of Data Mining Tools Using Case Study of Library Management System", International Journal of Information and Electronics Engineering, Vol. 3, No. 5, September 2013.

[11]. Ravinder Kumar, "A Convenient Hindi Language Interface to Database for Agriculture Based Queries", International Indexed & Referred Research Journal, ISSN-2250-2556, Nov-2014.

[12]. Rohini B. Kokare & Kirti H. Wanjale, "A Survey of Natural Language Query Builder Interface for Structured Databases using Dependency Parsing", International Journal of Computer Applications (0975 – 8887),Volume 107 – No 5, December 2014.

[13]. Poonam Gupta and Vishal Gupta(2013), "Algorithm for Punjabi Question Answering System", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7.