Database Interaction Using Automatic Speech Recognition

Pragati Naikare[1] Deeba Shahab[2] Shravani Deshmukh[3]
[1,2,3,]Department of Computer Engineering
[1,2,3]Dr. D. Y. Patil Institute of Engineering, Management & Research, India

Abstract— This Interaction with databases is achievable only if we are familiar about the standard SQL queries. In this paper, speech-based DBMS communication is the focus. Users can use their voice to interface with the database and retrieve information. Therefore, it is not necessary for the user to have prior understanding of SQL queries. Information retrieval is concerned with providing quick access to information based on a user's request, which will be presented as a query. If further information is required, a dialogue system that can understand spoken natural language questions will ask for it before responding to the speaker's question. Primarily a small number of information extraction research studies are devoted to the study of spoken language information extraction, the majority of which concentrate only on written language processing. In this system, user can also enter the query using speech. System will convert speech into the text format. This query will get transformed to SQL query. System will execute the query and gives output to the user.

Keywords: Natural Language Processing, LN2SQL, Speech Recognition, SQL, Syntactic, Semantic, Tokenization, Data Dictionary, Lexical Analysis

I. INTRODUCTION

Speech recognition technology has recently reached a higher level of performance and robustness, allowing it to communicate to another user by talking, and with the help of these it will recognize whole speech word by word. The isolated question word recognition based on the automated speech recognition system is applicable in speech-driven information retrieval and question answering system. The faculty or act of expressing or describing thoughts, feelings, or perceptions by the articulation of words. A form of communication in spoken query and text language, made by a speaker before an audience. Speech Recognition (SR) is the ability to translate a dictation or spoken query and word to text. Also known as “automatic speech recognition” (ASR), “computer speech recognition”, or “speech to query and text”.

This solution is intended for those who need to interact with databases but are not familiar with SQL. Speech recognition primarily concerns with a machine's or program's capacity to recognize spoken words and phrases and translate them into a machine-readable format. The software has the ability to recognize natural speech. The use of speech-to-text conversion will provide the answer to concerns with tokenization, syntactic, semantic, discourse, pragmatic analysis, and the utilization of dictionaries and grammars required for such analysis that arise in the analysis or generation of natural language text. Semantic Grammar is employed in the translation process.

The system will contain supporting features such as GUI and error handling module. On GUI, there is option ‘Click to Start’ to record speech. On the GUI, the user can see two things: spoken text and a converted SQL query for the same.

There are many challenges in the conversion of natural language query to SQL query like ambiguity which means that one word can have more than one meaning. In this case, one-word maps to more than one sense. Another challenge is the formation of complex SQL query and next challenge is about Discourse knowledge in which immediately preceding sentence affects the interpretation of next sentence for example if the user enters SELECT and INSERT query at the same time, then such a case is not understandable by the system

II. RELATED WORK

Prof. Debarati Ghosal, Tejas Waghmare, Vivek Satam, Chinmay Hajirnis proposed a system for "SQL query formation using natural language processing". This system prepares an expert system which converts simple natural language query into SQL query language by giving the user all possible intermediate queries so that user canselect appropriate intermediate query and the system will generate SQL query from intermediate one. Finally, system will execute the query and gives output to the user. [1].

Anum Iftikhar, Erum Iftikhar, Muhammad Khalid Mehmood proposed a system for "Domain Specific Query Generation from Natural Language Text". This system work for resolving ambiguity problem in NLP by using Stanford parser to parse English texts and generate SQL queries. The outcome of the suggested tool can be used for automated querying NoSQL databases and design from NL business rules. [2]

Prasun Kanti Ghosh, Saparja Dey, Subhabrata Sengupta proposed a system for "Automatic SQL Query Formation from Natural Language Query". This system will convert natural language query to SQL language and Speech to Text Recognition for Android using Python programming language. [3]

Prof. Sonal Gore, Niket Choudhary worked on “Impact of IntelliSense on the accuracy of Natural Language Interface to Database". This system work for an interface which allows a naive user to ask query to the database in his own language. The main purpose of building this system is 96% the highest accuracy is achieved and not all the concepts of SQL are considered to them. 4% error can also cause damage to a business. So this system proposes, to achieve 100% accuracy implementation of IntelliSense and an error handling module are involved in this system. Using machine learning, they enhanced IntelliSense so the machine could generate suggestions based on previously typed words. Those suggestions used to frame a complete and correct query. [4]

"Translating Controlled Natural Language Query into SQL Query using Pattern Matching Technique” is a system proposed by Rajender Kumar, Mohit Dua. This system work for allowing common users to give query to the databases and retrieve information from them using natural

language. A system is proposed as the user can give query in the Hindi language to retrieve data from the database. This system uses two analyzers such as morphological analyzer and word group analyzer. The main purpose of this analyzers is to extract the keyword from input Hindi query. Then for finding the type of keyword, it uses pattern matching technique. For reducing ambiguity, it uses the controlled Hindi language interface and suggested query feature. [5]

III. PROPOSED SYSTEM

First, the user has to login to the system. The user can ask queries, such as "give....", "put...", "I want...", "find...","search…". The user can enter the query in text format or in speech. The system will also provide facility to update and delete tables from the database. The user will also be able to insert data into the database. The query given by the user is processed step by step. There are four main steps or levels involved from Automatic Speech Recognition to SQL queries. These are as follows:

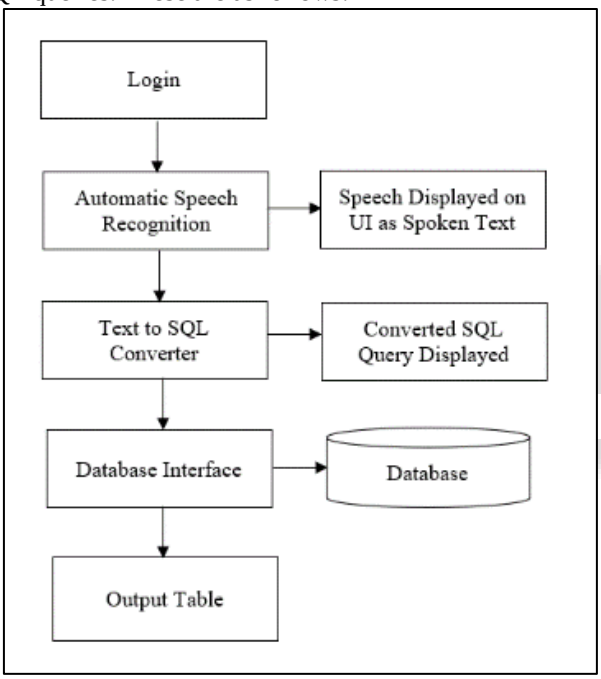


Fig. 1: System Architecture

A. Login

In order to add, edit, fetch or delete the data from the database, one must be an authorized person. To maintain the authenticity, we have a Login page which asks for the person's Employee id, Email id and Password. It's a password protected system. It will allow only if the password matches. If a person enters wrong password thrice then an alert would be sent to the admin, informing him about the action.

B. Automatic Speech recognition

Automatic Speech Recognition or ASR, as it's known in short, is the technology that allows human beings to use their voices to speak with a computer interface in a way that, in its most sophisticated variations, resembles normal human conversation.

The basic sequence of events that makes any Automatic Speech Recognition software, regardless of its sophistication, pick up and break down your words for analysis and response goes as follows:

1) Provide an audio feed to the device.
2) Speech is recorded as a wave file by the device.
3) The wave file is then cleaned by eliminating background noise and normalizing volume.
4) The filtered wave form is broken down into phonemes.
5) Each phoneme is like a chain link and by analysing them in sequence, starting from the first phoneme, the ASR software uses statistical probability analysis to deduce whole words and words to complete sentences
6) Your ASR, now having "understood" your words, can respond to you in a meaningful way.

The most advanced version of currently developed ASR technologies revolves around what is called Natural Language Processing, or NLP in short. NLP enables computers to understand natural language as humans do. Whether the language is spoken or written, natural language processing uses artificial intelligence to take real-world input, process it, and make sense of it in a way a computer can understand.

C. Text to SQL Converter

The output from Automatic Speech Recognition is to be converted into SQL query. This conversion takes place in four stages, namely, tokenization, lexical analysis, syntactic analysis, semantic analysis.

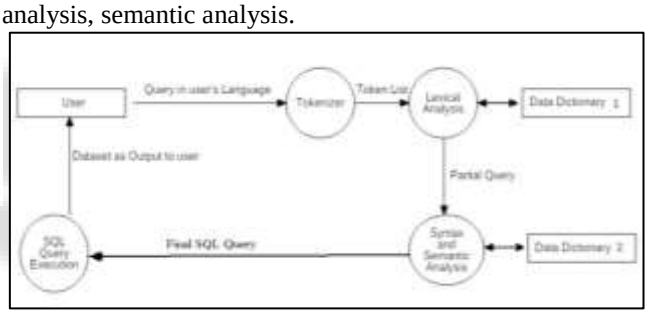


Fig. 2: Approach

1) Tokenization

The output from Automatic Speech Recognition is to be converted into SQL query. This conversion takes place in four stages, namely, tokenization, lexical analysis, syntactic analysis, semantic analysis. System will perform tokenization on the entered query by separating it into single words. Each word represents a token. Then these words will be stored in a separate list and passed to Lexical Analyzer.

2) Lexical Analysis

The tokenized list will be mapped with the dictionary. These words will get replaced by the database words from the dictionary and passed to syntactic analysis.

3) Syntactic Analysis

In this step dictionary of table names, attributes and keywords are maintained. Each tokenized word gets mapped with attributes in the dictionary. It is passed to Semantic Analysis for further processing.

4) Semantic Analysis

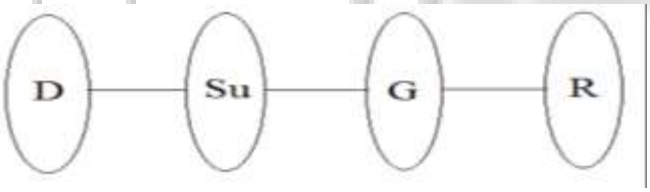
System will find words which represent conditions or symbols, and that word will get mapped with the dictionary. (For Example: If there is "less than or equal to" in the query, it will get mapped with the symbol "<=").

D. Database Interface

Python connect() method of mysql.connector module is used for database connectivity. We pass the database details like HostName, username, and the password in the method call, and then the method returns the connection object. Further the converted query is sent over this connection and the output is fetched in a table format.

Fig. 3: Output Table

IV. MATHEMATICAL MODULE



Where,
D= Document of user's voice input
Su = Speech recognition of the users
G = Construct query matching graph
R = Recommend query based on user's lifestyle
Set Theory

$$S= \{s, X, Y, e\}$$

Where,
S= Start of the program.
a) Register with system.
b) Login with web page.
X = Input of the program.
Users selected activity and lifestyles.
Y = Output of the program.
Recommend semantic based query to users.
e = End of the program.

A. Space Complexity:

The space complexity depends on Presentation and visualization of discovered patterns. More the storage of data more is the space complexity.

B. Time Complexity:

Check No. of patterns available in the datasets= n
If (n>1) then retrieving of information can be time consuming.
So, the time complexity of this algorithm is O(n).
= Failures and Success conditions.

C. Failures:

1) Huge database can lead to more time consumption to get the information
2) Hardware failure.
3) Software failure.

D. Success:

1) Search the required information from available in datasets.
2) User gets result very fast according to their needs.

V. CONCLUSIONS

The Database plays a vital role in business and industry. Considering the huge amount of data and its criticality the access of database should be limited to the admin who may or may not be a technical person. This approach will make the life easier for the admins who can provide access to people based on confidentiality, not on the basis of SQL knowledge. The system can convert speech in English, Hindi and German into SQL query with an accuracy of 85%. As our system is voice controllable it further adds on to the easiness with which data can be added, edited or deleted by the admins.

Future work and the scope of innovating would be initiated with adding numerous languages to the system so that a large population may utilize it. Also implementing discourse integration such that the system can recognize multiple SQL queries at once.

REFERENCES

[1] Prof. Debarati Ghosal, Tejas Waghmare, Vivek Satam, Chinmay Hajirnis "SQL query formation using natural language processing", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016.
[2] Prasun Kanti Ghosh, Saparja Dey, Subhabrata Sengupta "Automatic SQL Query Formation from Natural Language Query", International Journal of Computer Applications (0975 – 8887), International Conference on Microelectronics, Circuits and Systems (MICRO-2014).
[3] Biing-Hwang Juang, Fellow, IEEE, Wu Chou, Member, IEEE, and Chin-Hui Lee, Fellow" Minimum classification error rate methods for speech recognition, ‖IEEE Trans
[4] Abushariah,A.A.M,Gunawan,T.S,Khalifa,O..O,Abushariah, M.A.M. ―English digits speech recognition system based on hidden markov models‖ International conference on computer and communication engineering , 2010.
[5] Patten, T., & Jacobs, P. (1994). Natural-language processing. IEEE Expert, 9(1), 35. doi:10.1109/64.295134.
[6] "Natural Language Understanding" book by James Allen, Pearson Education, 2002.

[7] Huangi, Guiang Zangi, Phillip C-Y Sheu - A Natural Language database Interface based on the probabilistic context-free grammar, IEEE International Workshop on Semantic Computing and Systems 2008.
[8] Rajender Kumar, Mohit Dua "Translating Controlled Natural Language Query into SQL Query using Pattern Matching Technique", IEEE 2014.
[9] Anum Iftikhar, Erum Iftikhar, Muhammad Khalid Mehmood "Domain Specific Query Generation from Natural Language", Sixth International Conference on Innovative Computing Technology, IEEE2016.
[10] "Natural Language Understanding" book by James Allen, Pearson Education, 2002.

