

Automatic Whitelist Generation for SQL Queries Using Web Application Test

Venkati Mane¹, Jayesh Trivedi¹, Manalikamble¹, Shital Janjal¹, Prof. Vandana Chavan²

¹Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra India

²Professor Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra India

ABSTRACT

In this survey paper the proposed system will enable secure access of data to a voice- based user interface (UI) by enabling voice-based authentication and integration with an existing Natural Language Processing (NLP) system. This survey paper focuses on voice based SQL query generation. We study the question of how to improve the fetching the results from query results as well as applying the query to the database. Traditional predefined query forms are not able to satisfy various ad-hoc queries from users on those databases. Here, we propose Machine learning based technique to generate the SQL query based on user voice, a novel database query form interface, which is able to dynamically generate query forms.

Keywords: NLP, Languages and compilers, Optimization, Verification, Voice Recognition, Machine-independent microcode generation

I. INTRODUCTION

Natural Language Processing (NLP) is an area of application and research that explores how computers can be used to understand and manipulate natural language speech or text to do useful things. The foundation of NLP lie in a number of disciplines, namely, computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence robotics, and psychology. NLP researchers aim to gather knowledge on how human beings use and manipulate natural languages to perform desired tasks so that appropriate tools and techniques can be developed. Applications of NLP include a number of fields of study such as multilingual and cross-language information retrieval (CLIR), machine

transaction, natural language, text processing and summarization, user interfaces, speech recognition, artificial intelligence and expert systems.

While natural language may be the easiest system for people to learn and use, it has proved to be the hardest for a computer to understand. The goal of NLP is to enable communication between people and computers without resorting to memorization of complex commands and procedures.

In other words, NLP is a technique, which can make the computer understand the languages naturally used by humans. In this project, we are translating English query

Into a SQL query using semantic grammar. The system will accept users query in natural language as an input. The program will check whether the query is valid or not.

Then we will generate tokens by performing the division of the question clause. Each token represents a single word in the users query. The tokens from the query clause are compared with clauses already stored in the dictionary. The dictionary needs to be constantly updated. Then the algorithm scans the tokens and tries to find attributes present in the query. Then we find all the tables in the database which contain the attributes by comparing syntax and semantics. Then we build the final SQL query and execute it on the database and return the result dataset to the user.

II. PROBLEM STATEMENT

Most of used don't know about the data access, data searching, data query so, user have better knowledge that area. A voice controlled system for blind, which transceivers information in the form of audio: a personal virtual assistant which can take the human voice commands to perform tasks which otherwise would need the dependence on others.

III. LITERATURE SURVEY

Natural language processing can be done in two way communication with device one is written communication as well as verbal communication with device written communication is much more easier than the verbal communication. In written communication syntax, semantic, lexical and morphological analysis is done. Whereas in verbal communication includes all the process in written as well as additional process include additional knowledge about phonology as well as enough added information to handle the further ambiguities that arise in speech.

In this paper author represent a method for building a ldquonatural languages interfaces to data basesrdquo (NLIDB) system. [1].

system which is open source and can be accessed in remote locations using a simple hardware component. The processes including voice extraction, speech to text conversion, text processing and database management and speech synthesis have been explained in detail along with the python modules used to build the system. By minimizing the hardware components and using open source software, a universal, adaptable NLP system has been proposed [2].

The paper discusses natural language processing in field of Information retrieval. Natural Language Processing deals with human computer interaction. Natural language processing has a vital role in the field of Information retrieval. The storage and processing of information in response to user's query requires an information retrieval system that processes human or natural language. Various tasks of information retrieval such as Stemming, Lemmatization, POS tagging is discussed in the paper. The implementation of these tasks in Python is illustrated with examples.

The paper explains and implements natural language processing task in Information retrieval. Essential and preliminary tasks task required in almost all information retrieval system is explained with example in the paper. The tasks are performed in Python and with the inclusion of NLTK module. [3]. In this paper proposed system accepts user query in natural language and translate it into SQL query and retrieve result from database. Syntactic parsing, keyword extraction, stop words removal, cooccurrence matrix generation, use of WordNet, stemming algorithm and semantic mapping techniques have been used for formation of the SQL query from natural language input. Developed system gives correct answers of simple queries, queries with logical conditions and aggregate functions. As presented system does not support all forms of SQL queries, further development is necessary.

In this paper author discussed design and implementation of a system using modified word co-occurrence matrix method which will provide access to database using queries in English language [4].

In paper how to work are Recognizing speech using Sphinx, to voice base natural language processing using python [5]

We introduce a stochastic model for dialogue systems based on the Markov decision process. Within this framework we show that the problem of dialogue strategy design can be stated as an optimization problem, and solved by a variety of methods, including the reinforcement learning approach. The advantages of this new paradigm include objective evaluation of dialogue systems and their automatic design and adaptation. We show some preliminary results on learning a dialogue strategy for an air travel information system[6]

In this paper, we use information retrieval (IR) techniques to improve a speech recognition (ASR) system. The potential benefits include improved speed, accuracy, and scalability. Where conventional HMM-based speech recognition systems decode words directly, our IR-based system first decodes subword units. These are then mapped to a target word by the IR system. In this decoupled system, the IR serves as a lightweight, data-driven pronunciation model. Our proposed method is evaluated in the Windows Live Search for Mobile (WLS4M) task, and our best system has 12% fewer errors than a comparable HMM classifier. We show that even using an inexpensive IR weighting scheme (TF-IDF) yields a 3% relative error rate reduction while maintaining all of the advantages of the IR approach[7]

IV. PROPOSED SYSTEM

A. System Architecture:

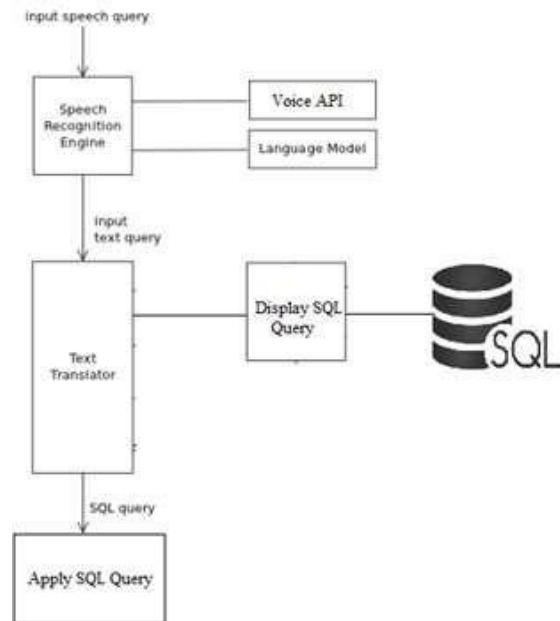


Fig 1. System architecture

This system has been under developed to enable secure access of data to a voice-based user interface (UI) by enabling voice-based authentication and integration with an existing Natural Language Processing (NLP) system.

We address the question of how to improve the fetching the results from query results.

Here, we propose Machine learning based technique to generate the SQL query based on user voice, a novel database query form interface, which is able to dynamically generate query forms.

B. Description:

1. In this system we implement the sql as a voice command and apply Natural Language Processing (NLP) with the help of Machine Learning technique.
2. First we use voice based input and after the input it is processed in the speech recognition engine.

In the speech recognition engine it has voice API and Language Model the 2

3. components of the speech recognition engine. Then the input text query is processed in the text translator and from the translator it is sent and then displayed in sql and then stored in the database that stored the sql query which will be applied in database. After the text translator query will be saved in the whitelist then the query will apply in the database.

V. APPLICATION

- This system scope is not limited. we use this system to the non-technical as well as technical user for understating the SQL Query.
- This kind of human-machine interfaces would allow a user to control a Database activity using the voice based.
- Government and private school.
- Database Management Company

VI. CONCLUSION

The implementation of the proposed system aims to translate SQL query into speech (voice). The scope of the project is to enhance the recognition capability for various SQL query and achieving more accuracy to generate the Database.

We address the question of how to improve the fetching the results from query results. Traditional predefined query forms are not able to satisfy various ad-hoc queries from users on those databases. Here, we propose Machine learning based technique to generate the SQL query based on user voice, a novel database query form interface, which is able to dynamically generate query forms.

VII. REFERENCES

- [1]. F.Siasar djahantighi, M.Norouzifard, S.H.Davarpanah, M.H.Shenassa, "USING NATURAL LANGUAGE PROCESSING IN ORDER TO CREATE SQL QUERIES", in Proceedings of the International Conference on Computer and Communication Engineering, Kuala Lumpur, May 2008, pp. 600-604.
- [2]. NATURAL LANGUAGE PROCESSING USING PYTHON International Journal of Scientific Engineering Research Volume 8, Issue 5, May-2017 19 ISSN 2229- 5518IEEE]
- [3]. Natural Language Processing Techniques Applied in Information Retrieval-Analysis and Implementation in Python, TulikaNarang, International Journal of Innovations Advancement in Computer Science IJIACS ISSN 2347 – 8616 Volume 5, Issue 4 April 2016
- [4]. Anuradha Mohite, Varunakshi Bhojane, "Natural Language Interface to Database Using Modified Co- occurrence Matrix Technique 2015 International Conference on Pervasive Computing (ICPC)
- [5]. Levin E., Pieraccini R., Eckert W. "Learning dialogue strategies within the Markov decision process framework", Automatic Speech Recognition and Understanding IEEE Proceedings, pp. 72-79, 1997.IEEE]
- [6]. Liddy, E. D. In Encyclopedia of Library and Information Science, 2nd Ed. Marcel Decker, Inc.
- [7]. J. Allan. Perspectives on information retrieval and speech. In Information Retrieval Techniques for Speech Applications: LNCS 2273, pages 1–10, 2002