International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

The Natural Language Query Using NLP by Generation of SQL Query

Prajkta Devgade[1], Jiwan Dehankar[2], Priyanka Bhende[3]

[1]Post Graduate Student, Department of CSE, Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India
[2]Professor, Department of CSE, Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India
[3]Professor, Department of CSE, Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India

---

Abstract - *Using command language for handling databases is usually knowledgeable and sophisticated problem. To overcome this, we proposed a model in NLP for converting the Natural Language Query to SQL query. This helps novice user to urge required content without knowing any complex details about SQL. A Natural Language Interface for Database acts as an intelligent interface for database which allows users to access the database by typing requests in their natural language. The steps involved during this process are tokenization, lemmatization, parts of speech tagging, parsing and mapping. The natural language interface is capable for translating the natural language query given by the user into an equivalent one in Database Query Language. The paper would give an overall view of the usage of Natural Language Processing (NLP) and use of normal expressions to map the query in English to SQL.*

***Key Words*: Natural language Processing, Structured Query language, Syntax semantic parsing, parsing.**

1. INTRODUCTION

Natural Language Processing may be a subfield of AI want to build intelligent computers which will interact with the person sort of a person. Most of the organizations use database to store different types of information like employee's data, financial data and statistical data related to company's growth, administrative information and so on. Database acts as the main source of information and organizes the data in a model that supports processes requiring information Main purpose of Natural Language Query Processing is for the interpretation of the English sentences by computer. The System proposes the architecture for processing the English Query fired by user to get SQL query using input as the text or speech. Speech recognition mainly deals with the ability of a machine or program to identify words and phrases in spoken language and convert them to a machine-readable format. Keyword based interface accepts query sort of program query which retrieves keyword from the input query and converts into SQL by applying rules from the generated knowledge domain. Natural Language Interface is applied on various formats of knowledge for natural language input query.

The system will contain supporting features such as GUI. On GUI, there the user can enter in natural language. When user will enter of in English Language, query will get converted to SQL query. The user can either view equivalent SQL query or can execute it directly. Objective of this paper is to convert a natural language query into a SQL to simplify data extraction.

2. LITERATURE REVIEW

Formation of SQL from Natural Language Query using NLP. Although several methodologies are employed to extract information from a database, Natural Language Processing has set a new standard in doing the same. This work presents a clear picture on the steps that are involved in NLP. Various processes like tokenization, lemmatization, syntactic and semantic analysis are administered to get the same SQL query from a natural language query. [1]

A review of various approaches in natural language interfaces to databases. This paper mainly attempts to introduce the basic concept of NLIDBs and different architectures of NLIDBs and we conclude the paper with some of the advantages and problems of using NLIDBs. [4]

Translating controlled natural language query into SQL query using pattern matching technique. In this paper, we present an approach to convert a Hindi sentence into SQL query. We implement a system that uses the morphological analyzer and word group analyzer to extract and tag the keyword. Domain-oriented dictionary are used to determine the type of keyword used in input query. [8]

Conversion of Natural Language Query to SQL Query. Use of Natural Language brings ease for any person. This system will help T&P officer to easily retrieve and manage data from student database using simple English language. There is no need for the user to learn complex query language like SQL. [3]

Using natural language processing so as to make SQL queries. By providing an expert system, we are encoding hidden mystery of natural language; the very fact that common words tend to possess multiple meanings can cause ambiguity, the expert system can maintain database that represents the state of the world by looking at the context surrounding the sentences and receives the best recognized from the text. [11]

3. PROPOSED SYSTEM

The proposed system is meant to attenuate the communication gap between a person's and computer. A system is meant which contains an intelligent layer that accepts common user's sentences in natural language as input, converts these sentences into standard SQL queries and executes them to retrieve data from relational databases. Our proposed system consists of several modules that are used to extract key words alone and leave out the redundant data. This is critical because presence of redundant data will certainly decrease the overall performance of the system. Input data initially goes through an NLP phase followed by a mapping phase.

3.1 Graphical User Interface (GUI)

The system has a user-friendly GUI which consists of a query editor through which the user can provide the English language query to the system and a query generator that provides the resultant Query of the user's request.

3.2 Natural Language Processor (NLP)

The input provided by the user must be processed to identify the constraints and predicates required to formulate the desired SQL query. The NLP makes use of techniques such as tokenization, part of speech tagging, chunking and entity recognition. The NLP provides chunked tags which are processed to analyses the attributes and predicates such as table name, attribute name or column names and selection conditions or criteria for the input user's query.

3.3. Query Generator

The query generator makes use of a query translation. The algorithm helps in identification of predicates based on the set of rules. A prediction model is, used when the translation algorithm fails to generate sufficient information to formulate the query, also included in the query generator to predict the query. The formulation of query is done with the help of pre-defined structure and attributes and predicates identified using NLP.

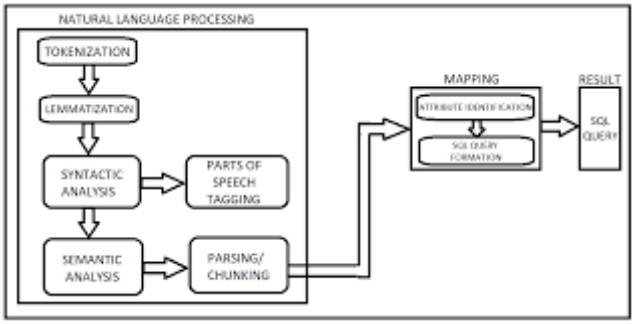

Fig -1: Proposed System Model

In the pre-processing phase, the input sentence is split into tokens. From these tokens, root is extracted by getting obviate inflectional morphemes. This is done with the help of lexicon and suffix replacement rules. Second phase of database component is database management system (DBMS) which execute the SQL query over the database.

Third phase of database component is response generator. It takes the output of DBMS as its input and converts it into the English language.

4. CONCLUSIONS

In this paper, we present an approach to convert an English sentence into SQL query. the result of using an expert system beside common existing solutions for transforming natural language expressions to SQL query language. There is no need for

the user to learn complex query language like SQL. The facility to accept the input in speech format makes the system user-friendly. Our system will convert natural language query into SQL language query and provides required information.

REFERENCES

[1] Uma, M., Sneha, V., Sneha, G., Bhuvana, J., & Bharathi, B. (2019). Formation of SQL from Natural Language Query using NLP. 2019 International Conference on Computational Intelligence in Data Science (ICCIDS).

[2] Das, A., & Balabantaray, R. C. (2019). MyNLIDB: A Natural Language Interface to Database. 2019 International Conference on Information Technology (ICIT).

[3] Kate, A., Kamble, S., Bodkhe, A., & Joshi, M. (2018). Conversion of Natural Language Query to SQL Query. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA).

[4] Reshma, E. U., & Remya, P. C. (2017). A review of different approaches in natural language interfaces to databases. 2017 International Conference on Intelligent Sustainable Systems (ICISS).

[5] Li, Q., Li, L., Li, Q., & Zhong, J. (2020). A Comprehensive Exploration on Spider with Fuzzy Decision Text-to-SQL Model. IEEE Transactions on Industrial Informatics, 16(4), 2542–2550.

[6] Mahmud, T., Azharul Hasan, K. M., Ahmed, M., & Thwoi Hla Ching Chak. (2015). A rule-based approach for NLP based query processing. 2015 2nd International Conference on Electrical Information and Communication Technologies (EICT).

[7] Sangeeth, N., & Rejimoan, R. (2015). An intelligent system for information extraction from relational database using HMM. 2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI).

[8] Kumar, R., & Dua, M. (2014). Translating controlled natural language query into SQL query using pattern matching technique. International Conference for Convergence for Technology-2014.

[9] Agrawal, R., Chakkarwar, A., Choudhary, P., Jogalekar, U. A., & Kulkarni, D. H. (2014). DBIQS — An intelligent system for querying and mining databases using NLP. 2014 International Conference on Information Systems and Computer Networks (ISCON).

[10] Shah, A., Pareek, J., Patel, H., & Panchal, N. (2013). NLKBIDB - Natural language and keyword-based interface to database. 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI).

[11] Siasar djahantighi, F., Norouzifard, M., Davarpanah, S. H., & Shenassa, M. H. (2008). Using natural language processing in order to create SQL queries. 2008 International Conference on Computer and Communication Engineering.

[12] Adwait Ratnaparkhi, "Learning to Parse Natural Language with Maximum Entropy Models," Machine Learning, vol. 34, issue 1-3, Feb. 1999, pp. 151-175.0

[13] I. Androutsopoulos, G. D. Ritchie and P. Thanisch, “Natural Language Interfaces to Databases – An Introduction,” in Natural Language Engineering, vol. 1, part 1, 1995, pp. 29-81

[14] George A. Miller (1995), “WordNet: A Lexical Database for English,” in Communications of the ACM Vol. 38, No. 11: 39-41.

[15] Ana-Maria Popescu, Oren Etzioni and Henry Kautz, “Toward a Theory of Natural Language Interfaces to Databases,” in IUI Proceedings 2003.