AI model to generate SQL queries from natural language instructions through voice

Presented By:

Anuja Patil-(501848)

Rohit Raina-(501855)

Aditya Sawant-(501860)

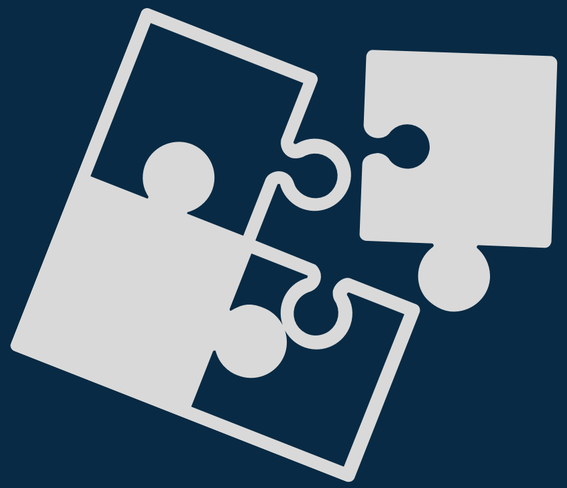
Guided By:

Prof.Anand Pardeshi

Abstract:

- The proposed model describes the generation of Structured Query Language(SQL) from natural language input, spoken(audio input) by the user.
- The model is constructed on NLP (Natural Language Processing) and Neural Networks (Deep Learning) technologies. Model is trained on the dataset with natural language as input and the outline skeletal structure of the query as output.
- Then the output will be processed and the final output will be displayed to the user. The project aims to benefit the people who are suffering from Repetitive Stress Injury (RSI), which has been attributed to work requiring a long period of typing and also to those who are not familiar with SQL queries. As this system will readily provide the required query.

Problem Statement



- The structured query language is the most used tool for managing data in the relational database system. Most users are not familiar with SQL which is problematic for managing and retrieving data.

- Though the user is good at coding, but every coder makes some error and it take time to find and resolve the error mainly syntactical error. And many a time, mainly coder faces an issue of repetitive stress injury in fingers due to constant movement of fingers while typing the code, So they have to go through a lot of pain while typing the code.

Literature Survey

Title	Author	Description
Automatic SQL Query Formation from Natural Language Query (2014)	Pranali Nagare, Smita Indhe, Dhanashri Sabale	This paper proposes a system which divided input in tokens and then generated the output after processing.
Extracting Sql Query Using Natural Language Processing (2019)	Nandhini S, B.Viruthika, Almas Saba, Suman Sangeeta Das	This paper describes the model which uses lexical analysis, syntax analysis and semantic analysis for query generation.
Natural Language Processing with some Abbreviation to SQL (2018)	Chandrakala Kombade, Monika More, Shweta Patil, Anjali Devi Pujari	They developed a system to generate valid SQL queries after parsing natural language using open source tools and libraries.

Title	Author	Description
A Model of a Generic Natural Language Interface for Querying Database (2018)	Bais Hanane and Mustaph Machkour	In this paper they proposed model which predicts the query based on given input and the ability to improve knowledge based on machine learning approach
Automatic SQL Query Formation from Natural Language Query (2014)	Prasun Kanti Ghosh, Saparja Dey, Subhabrata Sengupta	This paper describes the model which uses lexical analysis, syntax analysis and semantic analysis for query generation.

Proposed System

In the proposed system, the user has to speak his requirement into the application and then the speech will be converted to text. The text will further be processed and send to model has input. The model will return the executable query.

Eg:

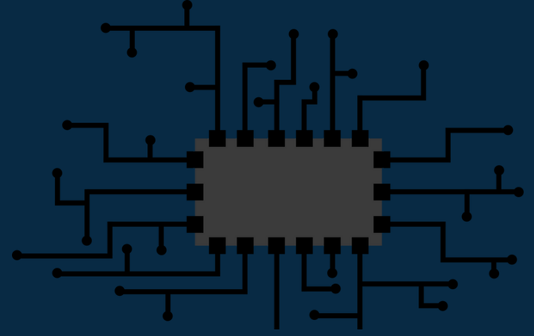
Input: Give me the marks column from the student table.

Output: select marks from student.

Input: give me name of students whose marks are greater than 95.

Output: select name from student where marks >95.

Proposed System



The user will speak his requirements to the application

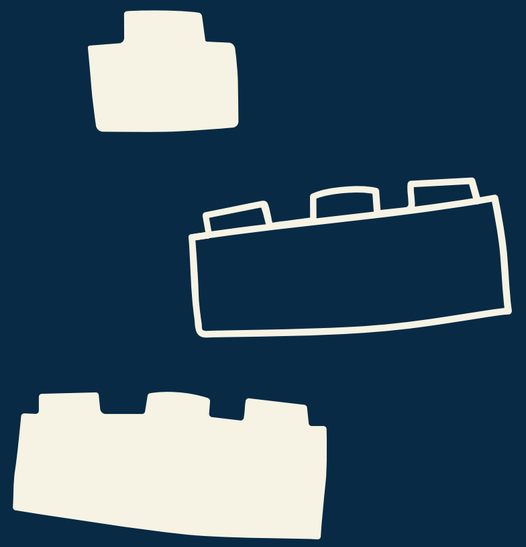
Speech will be converted to text

The text will be processed, and brought into required format

The processed text will be send to model

The model will predict , and processed the output and return query

Model Building Steps:



01 Creation of dataset.

02 Pre-processing of dataset

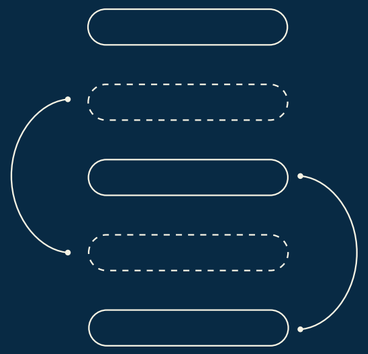
03 Building LSTM Model

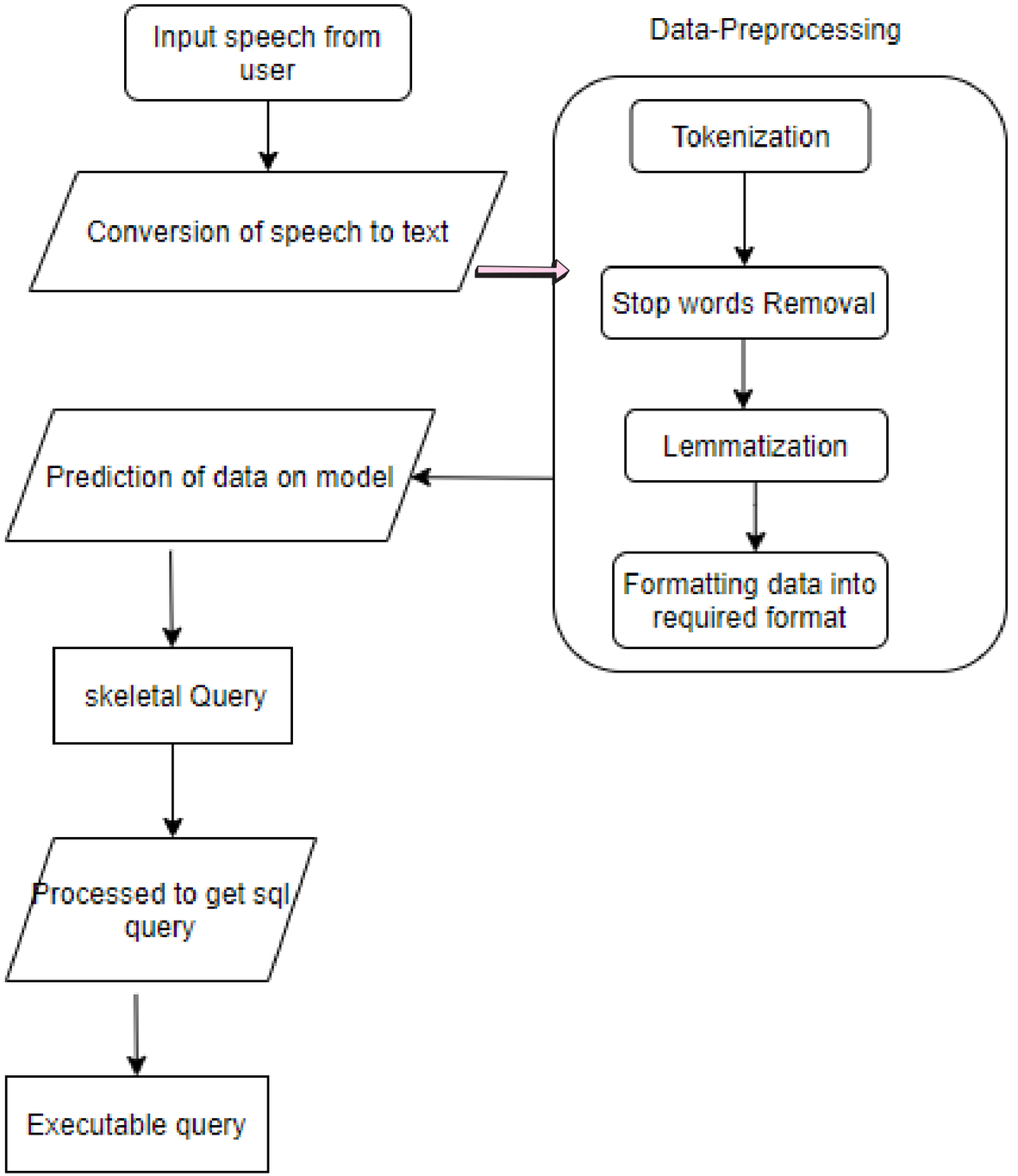
04 Training Model on dataset

05 Fetching Result

Flowchart

It depicts the entire flow of process from the user giving inputs and getting an executable query.





Future Scope

In this project we are implementing the system which is currently capable of handling simple queries along with some complex queries. Because all the forms of SQL queries are not supported, further development would be required.

In future we plan to implement a fully functional platform where the user or programmer will be able to code using voice commands only. The additional functionality that can be developed is switching between the languages in which the user wants to code. So the need of knowing syntax of different languages is eliminated.

Conclusions 

- Natural Language Processing is very powerful tool which can change the complete working of the computer program interface. The aim is to evaluate correct SQL queries by using NLP.
- Use of Natural Language helps user to easily retrieve data. This system will help many organizations such as education, medical,etc. to easily retrieve and manage data from database using simple English language.
- There is no need for the user to learn complex query language like SQL. The facility to accept the input in speech as well as in text format makes the system user-friendly. Our system will convert natural language query into SQL language query and provides required information from database to organizations.

References

[1]Automatic SQL Query Formation from Natural Language Query by Pranali Nagare, Smita Indhe, Dhanashri Sabale

[2]Extracting Sql Query Using Natural Language Processing by Nandhini S, B.Viruthika, Almas Saba, Suman Sangeeta Das

[3]Natural Language Processing with some Abbreviation to SQL by Chandrakala Kombade, Monika More, Shweta Patil, Anjalidevi Pujari

[4]A Model of a Generic Natural Language Interface for Querying Database by Bais Hanane and Mustaph Machkour

[5]Automatic SQL Query Formation from Natural Language Query by Prasun Kanti Ghosh, Saparja Dey, Subhabrata Sengupta

Thank you!