



Dissertation on

“Voice Assistant For PESU”

Submitted in partial fulfilment of the requirements for the award of degree of

Bachelor of Technology

in

Computer Science & Engineering

UE20CS390A – Capstone Project Phase - 1

Submitted by:

Ankit Borker	PES1UG21CS805
Arya Girigoudar	PES1UG21CS806
Nitin Gaurav	PES1UG21CS826
Vishnu Patgar	PES1UG21CS844

Under the guidance of

Prof. Priya Badrinath
Assistant Professor

January - May 2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

FACULTY OF ENGINEERING

PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)

100ft Ring Road, Bengaluru – 560 085, Karnataka, India



PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)

100 Feet Ring Road, Bengaluru – 560 085, Karnataka, India

FACULTY OF ENGINEERING

CERTIFICATE

This is to certify that the dissertation entitled

‘Voice Assistant for PESU’

is a bonafide work carried out by

Ankit Borker
Arya Girigoudar
Nitin Gaurav
Vishnu Patgar

PES1UG21CS805
PES1UG21CS806
PES1UG21CS826
PES1UG21CS844

in partial fulfilment for the completion of sixth semester Capstone Project Phase - 1 (UE19CS390A) in the Program of Study - **Bachelor of Technology in Computer Science and Engineering** under rules and regulations of PES University, Bengaluru during the period Jan. 2022 – May. 2022. It is certified that all corrections / suggestions indicated for internal assessment have been incorporated in the report. The dissertation has been approved as it satisfies the 6th semester academic requirements in respect of project work.

Signature
Priya Badrinath
Designation

Signature
Dr. Shylaja S S
Chairperson

Signature
Dr. B K Keshavan
Dean of Faculty

External
Viva

Name of the Examiners

Signature with Date

1. _____

2. _____

DECLARATION

We hereby declare that the Capstone Project Phase - 1 entitled **“Voice Assistant For PESU”** has been carried out by us under the guidance of Prof. Priya Badrinath, Assistant Professor and submitted in partial fulfilment of the completion of sixth semester of **Bachelor of Technology in Computer Science and Engineering** of **PES University, Bengaluru** during the academic semester January – May 2022. The matter embodied in this report has not been submitted to any other university or institution for the award of any degree.

PES1UG21CS805

Ankit Borker

PES1UG21CS806

Arya Girigoudar

PES1UG21CS825

Nitin Gaurav

PES1UG21CS844

Vishnu Patgar

ACKNOWLEDGEMENT

We would like to express our gratitude to Prof. Priya Badrinath, Department of Computer Science and Engineering, PES University, for her/his continuous guidance, assistance, and encouragement throughout the development of this UE19CS390A - Capstone Project Phase – 1.

We are grateful to the project coordinator, Dr. Priyanka H., all the panel members & the supporting staff for organizing, managing, and helping the entire process.

We take this opportunity to thank Dr. Shylaja S S, Chairperson, Department of Computer Science and Engineering, PES University, for all the knowledge and support we have received from her.

We are grateful to Dr. M. R. Doreswamy, Chancellor, PES University, Prof. Jawahar Doreswamy, Pro Chancellor – PES University, Dr. Suryaprasad J, Vice-Chancellor, Dr. B.K. Keshavan, Dean of Faculty, PES University for providing us various opportunities and enlightenment during every step of the way.

Finally, this project could not have been completed without the continual support and encouragement we have received from our family and friends.

ABSTRACT

The purpose of this project is to create an AI enabled voice assistant for PESU which will allow students to get their information via voice commands, this is going to make PESU access easier and user friendly, as they will be able to get MARKS, ATTENDANCE, COURSES INFO and RESULTS easily using voice commands only.

The scope of a PESU voice assistant is vast and continues to expand as new use cases and functionalities are developed. The potential for voice assistants is limited.

1. Easier access to PESU resources
2. Get the latest information about academics

Table of Content

1. Introduction 3
2. Problem statement 4-5
3. Literature Review 6-10
 - 3.1.1 Data Extraction
 - 3.1.2 Reduction of Noise From Voice:
 - 3.1.3 DATABASE HANDLING:
 - 3.1.4 Security component:
4. Data 11-12
 - 4.1 Overview
 - 4.2 Types of data required for the system
 - 4.2.1 Input Data
 - 4.2.2 Model Architecture
 - 4.2.3 Training Details
 - 4.2.4 Performance Metrics

5. SRS 13-17

5.1 Product Perspective

5.1.1 Product Features

5.1.2 Operating Environment

5.2 Functional Requirements

5.3 External Interface Requirements

5.3.1 Hardware Requirements

5.3.2 Software Requirements

5.4 Non-Functional Requirements

5.4.1 Performance Requirement

5.4.2 Compatibility Requirements

5.4.3 Usability Requirements

6. System Design 18-28

6.1 Design Considerations

6.1.1 Design Goals

6.1.2 Architecture Choices

6.1.3 High Level Design

6.1.4 Use Case Diagram

6.1.5 Swim lane Diagram

6.3.3 System design architecture for voice assistant:

6.3.4 Flow chart for voice assistant working:

6.3.5 User Interface Diagrams

6.1.3 Constraints, Assumptions and Dependencies

7. Capstone Phase – I 29-31

7.1 CONCLUSION

8. PLAN OF WORK FOR CAPSTONE PROJECT PHASE-2

INTRODUCTION

1.INTRODUCTION

The purpose of this project is to create an AI enabled voice assistant for PESU which will allow students to get their information via voice commands, this is going to make PESU access easier and user friendly, as they will be able to get MARKS, ATTENDANCE, COURSES INFO and RESUTLS easily using voice commands only.

The requirements for voice assistant are students can get their information without going into website and just by talking to the assistant, which will fetch their data directly from the database without going in depths of the website.

The product is a PESU of software which is going to:

- Speech recognition: The assistant will have ability to understand and interpret spoken language.
- Natural language processing: The assistant will have ability to understand and interpret human language, including the nuances of grammar and context.
- Text-to-speech synthesis: The assistant will have ability to convert written text into spoken language.
- Voice commands: The assistant will have ability to respond to spoken commands and perform tasks based on those commands.
- Personalization: The ability to recognize individual users and provide personalized responses based on their preferences and past behaviour.

PROBLEM STATEMENT

2.PROBLEM STATEMENT

Voice assistants can help individuals manage their personal lives by setting reminders, scheduling appointments, showing results, reminding about attendance and read notifications to students. The scope of a PESU voice assistant is vast and continues to expand as new use cases and functionalities are developed. The potential for voice assistants is limited only by the imagination of developers and the needs of users.

PRODUCT PERSPECTIVE:

Integration: PESU voice assistant will be seamlessly integrated into the product ecosystem, whether it is a standalone device or part of a larger system.

Compatibility: PESU voice assistant will be designed to work with a wide range of devices and platforms, including smartphones, tablets, smart speakers, and other connected devices.

User interface: The user interface of our PESU voice assistant will be intuitive and easy to use, allowing users to interact with the device using natural language commands.

Functionality: PESU voice assistant will be designed to perform a wide range of tasks, from simple tasks like setting reminders and playing music to more complex tasks.

Performance: PESU voice assistant will be designed to provide fast, accurate, and reliable performance, even in complex and dynamic environments

LITERATURE REVIEW

3.LITERATURE REVIEW

3,1DATA EXTRACTION: The data extraction component will help in extraction of data from pdf format, image documents to extract data from image has to be in good quality. That's why we are using Logistic Regression to classify image as good or bad.

- For Extraction of data from. We are going to use OpenCV and PY tesseract for text extraction from each cell.
- Then Cells are detected and text are extracted and CSV/JSON is generated
- Logistic regression is used for checking quality of image
- Performance Is optimized by creating a background by calculating the height and width from all the cell by ROI crops

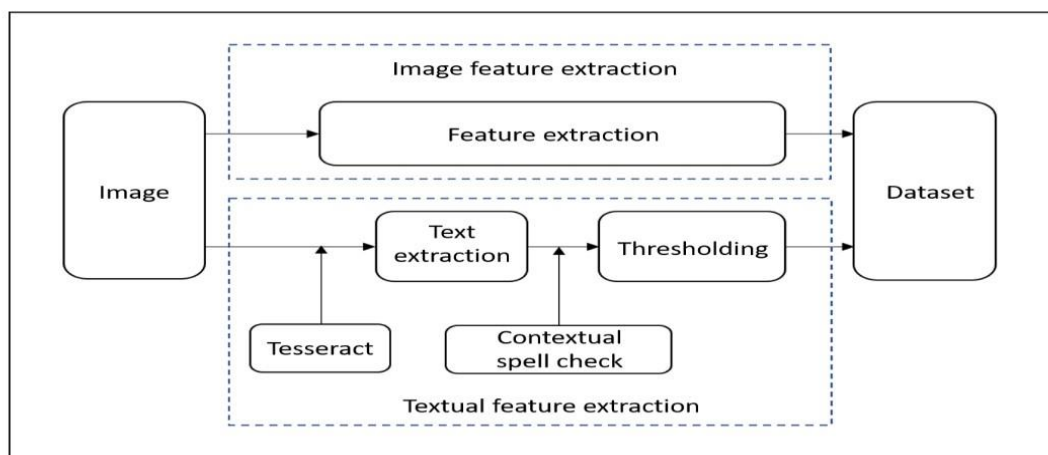


Fig. 3: Schematic flow of dataset compilation.

3.2 Reduction of Noise From Voice:

Pre-processing of voice is required to remove noise

Working:

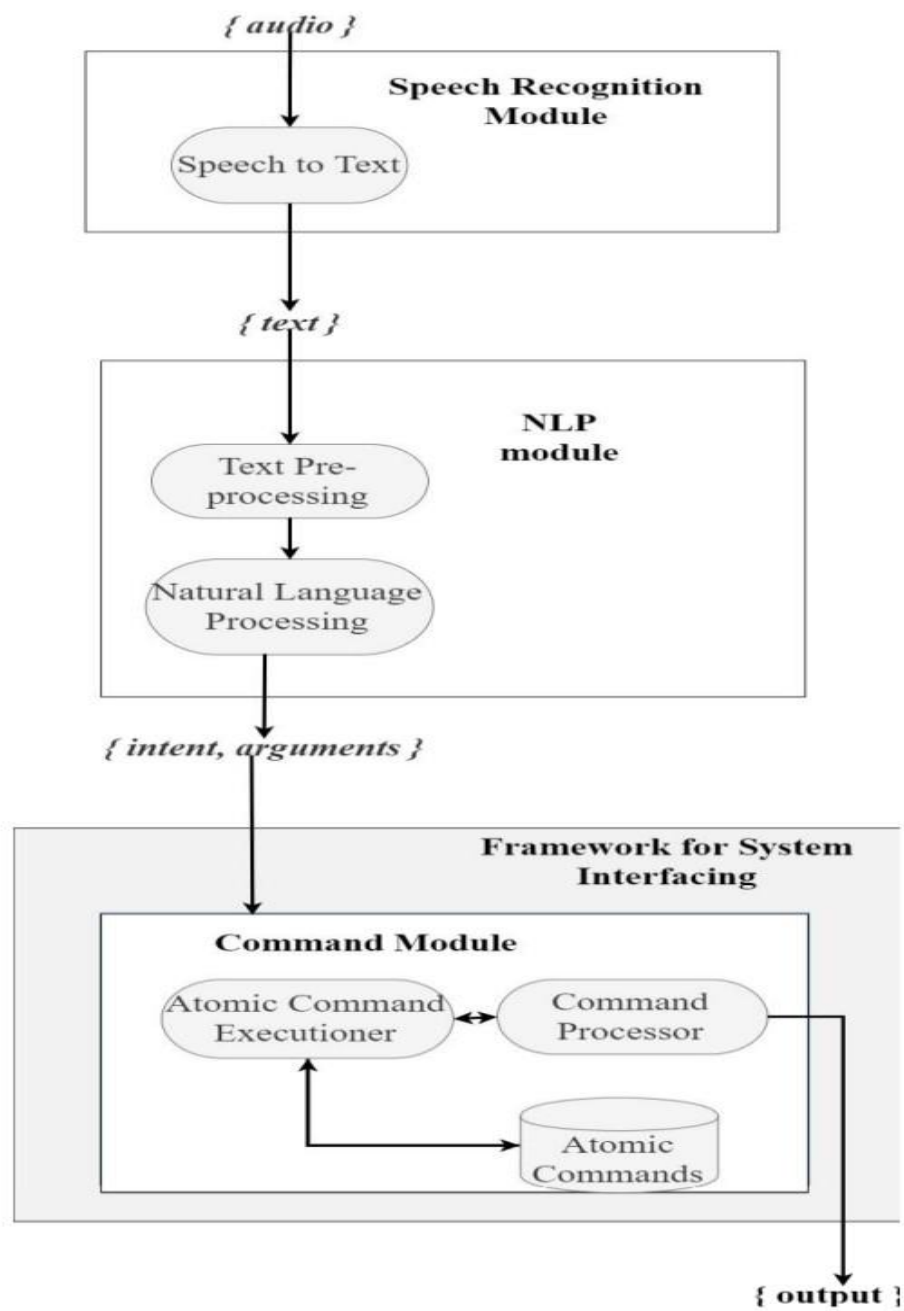
- Audio is broken down into 0.125 seconds clip, fed to feature extraction algorithm
- Then, weights are added and passed to KNN, KNN classifies as human voice and non-human voice

Audio Features: -

- Mel-frequency Cepstral Coefficients
- Spectral Roll – Off
- Spectral Centroid
- Zero Crossing Rate

The Speech Recognizing library will be used in 2 variants:

- online and offline modes, only difference is in online mode audio is passed to Certain API's
- on the other hand, offline mode makes use of the popular “CMU Sphinx” library for conversion of speech to text.



3.3 DATABASE HANDLING:

The data will be fetched directly from the PESU dummy database (which we are going to create then we'll put into production) , via queries generated by the assistant

Working:

- The assistant will generate queries from the voice model and every word will be tokenized as verbs and nouns
- The words will then be checked from selected key-words to fetch the information from the database

3.3 Security component:

- Since the voice assistant will handle lots of student's data and handling and managing of confidential data is a risk factor
- So different encrypting algorithms will be used as means of keeping data secure from the intruders
- The voice assistant model will keep the data secure within database and itself

DATA

4.DATA

4.1 Overview

We need to Capture the voice of student and sent it to server and Server will do the following steps

- Pre-processing Voice to remove noise from it.
- Convert it to text.
- Convert it into query, execute and store the result into database. Cache if necessary

4.2 Types of data required for the system

4.2.1 Input Data

- We mainly Require Data of the students in some kind of format.
- And Voice Command of Students. The size depends mainly on the Command and No. of Queries (ex: - Get me my latest GPA and What were my ISA Marks and their average)

4.2.2 Model Architecture

- **Tokenization:** - Spacy, NLTK.
- **PESU:** - HMM Model
- **Named Entity:** - CRF (Conditional Random Field)

4.2.3 Training Details

- **We're Going to use Pre-trained Models.**

4.2.4 Performance Metrics

- Accuracy of models are roughly 84% of all models

SYSTEM REQUIREMENTS SPECIFICATION

5. SYSTEM REQUIREMENTS SPECIFICATION

5.1 Product Perspective

Integration: PESU voice assistant will be seamlessly integrated into the product ecosystem, whether it is a standalone device or part of a larger system.

Compatibility: PESU voice assistant will be designed to work with a wide range of devices and platforms, including smartphones, tablets, smart speakers, and other connected devices.

User interface: The user interface of our PESU voice assistant will be intuitive and easy to use, allowing users to interact with the device using natural language commands.

Functionality: PESU voice assistant will be designed to perform a wide range of tasks, from simple tasks like setting reminders and playing music to more complex tasks.

Performance: PESU voice assistant will be designed to provide fast, accurate, and reliable performance, even in complex and dynamic environments.

5.1.1 Product Features

- **Speech recognition:** The assistant will have ability to understand and interpret spoken language.
- **Natural language processing:** The assistant will have ability to understand and interpret human language, including the nuances of grammar and context.
- **Text-to-speech synthesis:** The assistant will have ability to convert written text into spoken language.
- **Voice commands:** The assistant will have ability to respond to spoken commands and perform tasks based on those commands.
- **Personalization:** The ability to recognize individual users and provide personalized responses based on their preferences and past behaviour.

5.1.2 Operating Environment

1. Power source: Many voice assistants are designed to be always-on and require a stable power source to operate reliably.
2. Internet connectivity: A voice assistant needs a reliable and stable internet connection to function properly.
3. Compatibility with devices: Voice assistants are designed to work with various devices such as smartphones, smart speakers.
4. Access to data: To function effectively, a voice assistant needs access to data such as user preferences, history, and other relevant information.

5.2 Functional Requirements

- Voice Authorization
- Generating Intents and Arguments Then sending Arguments to database Caching most asked Queries
- NLP module for extracting intents and arguments
- Contextual awareness

5.3 External Interface Requirements

1. The project will be GUI based and will also support voice commands.
2. It will majorly be Voice based commands/Inputs

5.3.1 Hardware Requirements

- Any device with internet connection and has a screen with working microphone can use the voice assistant.

5.3.2 Software Requirements

- None

5.4 Non-Functional Requirements

- Security for student's Data
- It must have user friendly UI
- Must Be Accurate with the information it provides
- And Must authorize Correct user according to their SRN

5.4.1 Performance Requirement

- Response Time: The voice interface must respond to user requests within a specific time frame, usually measured in seconds or milliseconds.
- Throughput: The voice assistant must be able to handle a certain number of requests or transactions per second, depending on the expected usage.
- Availability: The assistant must be available for use during certain hours of the day, with a defined percentage of uptime (e.g., 99.9% uptime).
- Scalability: The project must be able to handle an increasing number of users or requests without sacrificing performance or reliability.
- Resource utilization: The project must use system resources efficiently, such as CPU, memory, and network bandwidth, to avoid performance degradation or system crashes.

5.4.2 Compatibility Requirements

- **Security and Data Integrity:** The assistant must be secure, with measures in place to prevent data breaches or unauthorized access, and must ensure the integrity of data.
- **Maintenance and Support:** The assistant must be easy to maintain, and support should be readily available in case of issues or bugs.

5.4.3 Usability Requirements

None, as it will be easy to use via voice commands only.

SYSTEM DESIGN

6.SYSTEM DESIGN

6.1 Design Considerations

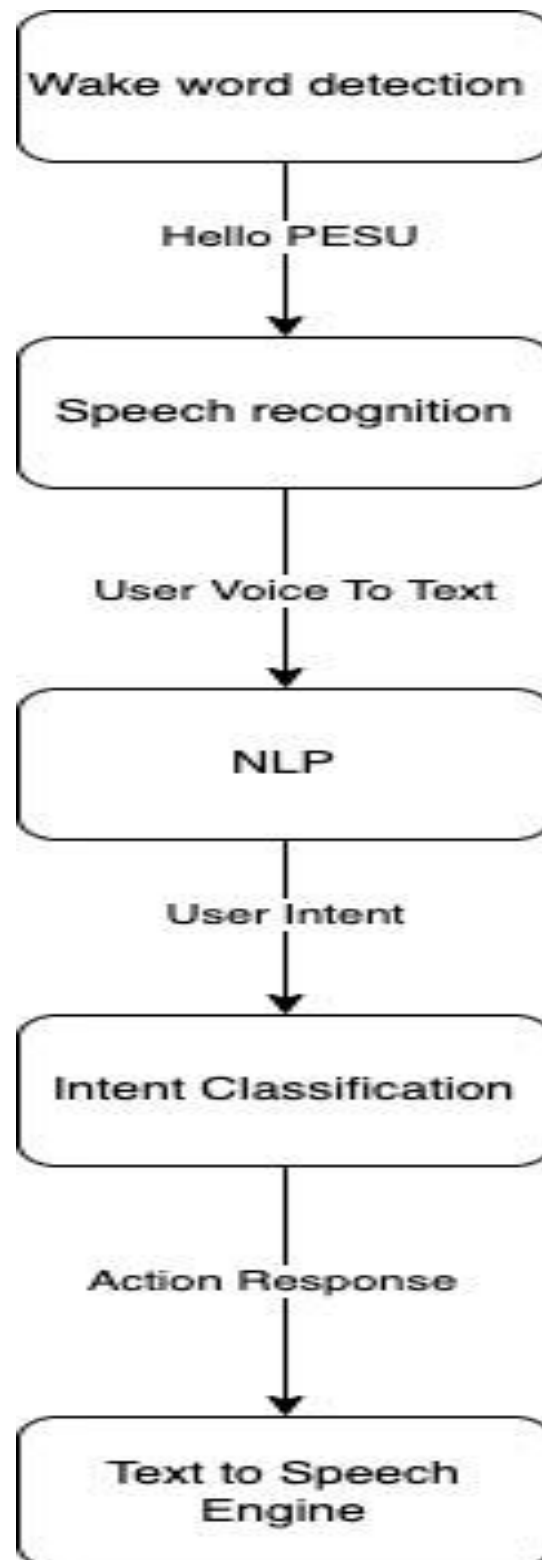
6.1.1 Design Goals

The main goal of designing a project is to describe the working process which will ease the work of developer. We are giving an analogy of how the components of product will be working.

Although the following model is predicated on one, it is not actually based on a client-server architecture.

6.1.2 Architecture Choices

We have Several choices when it comes to architecture, as voice assistant are booming these days. But we have decided to go with simplest approach below is diagram of our choice



6.1.3 Constraints, Assumptions and Dependencies

Assumptions made

- PESU will allow us to use Their Website/database.
- This is only limited to Students only

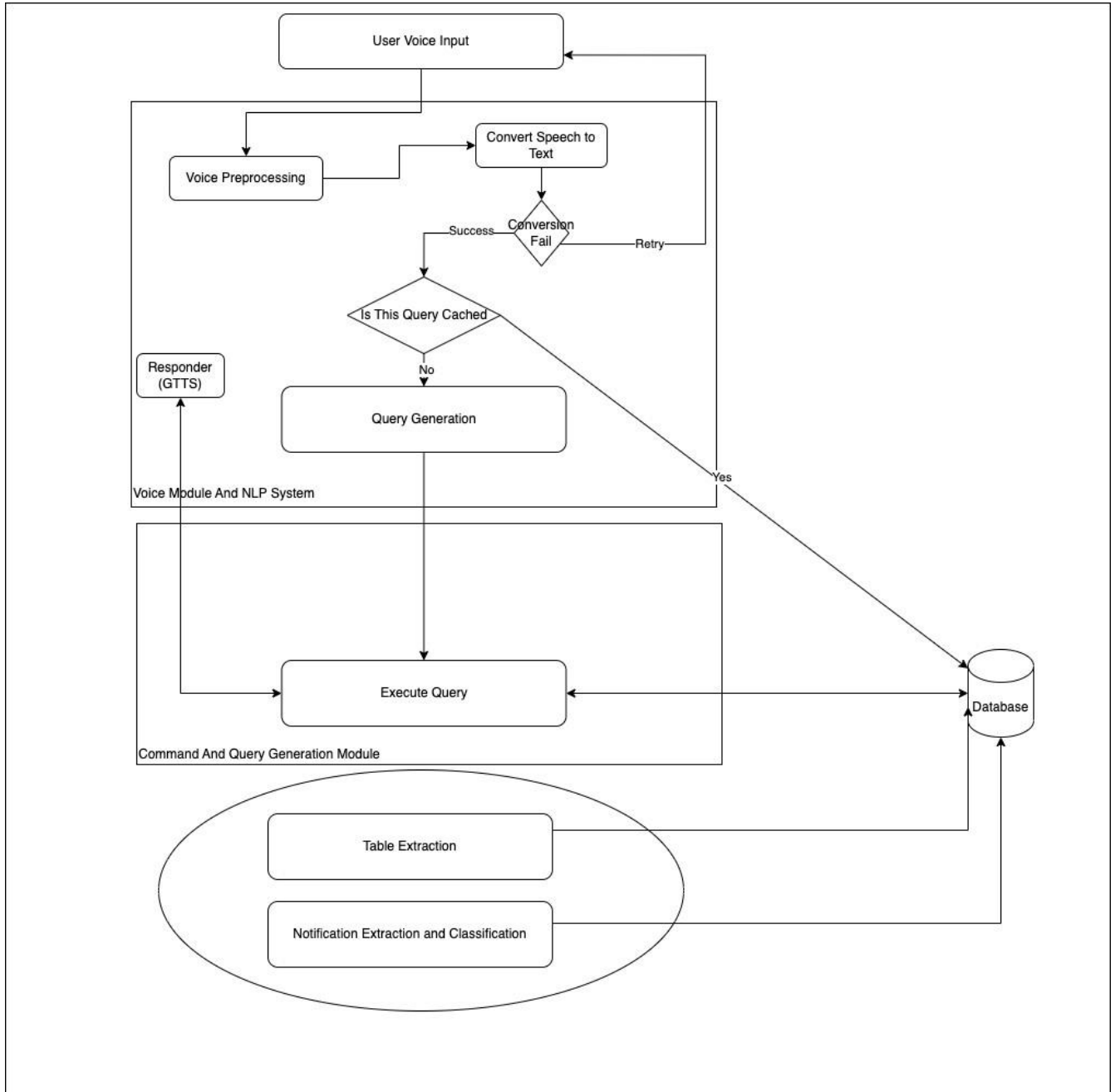
Risk involved

- Our only risk is of leakage of confidential data of students. We need Maximum Security for this

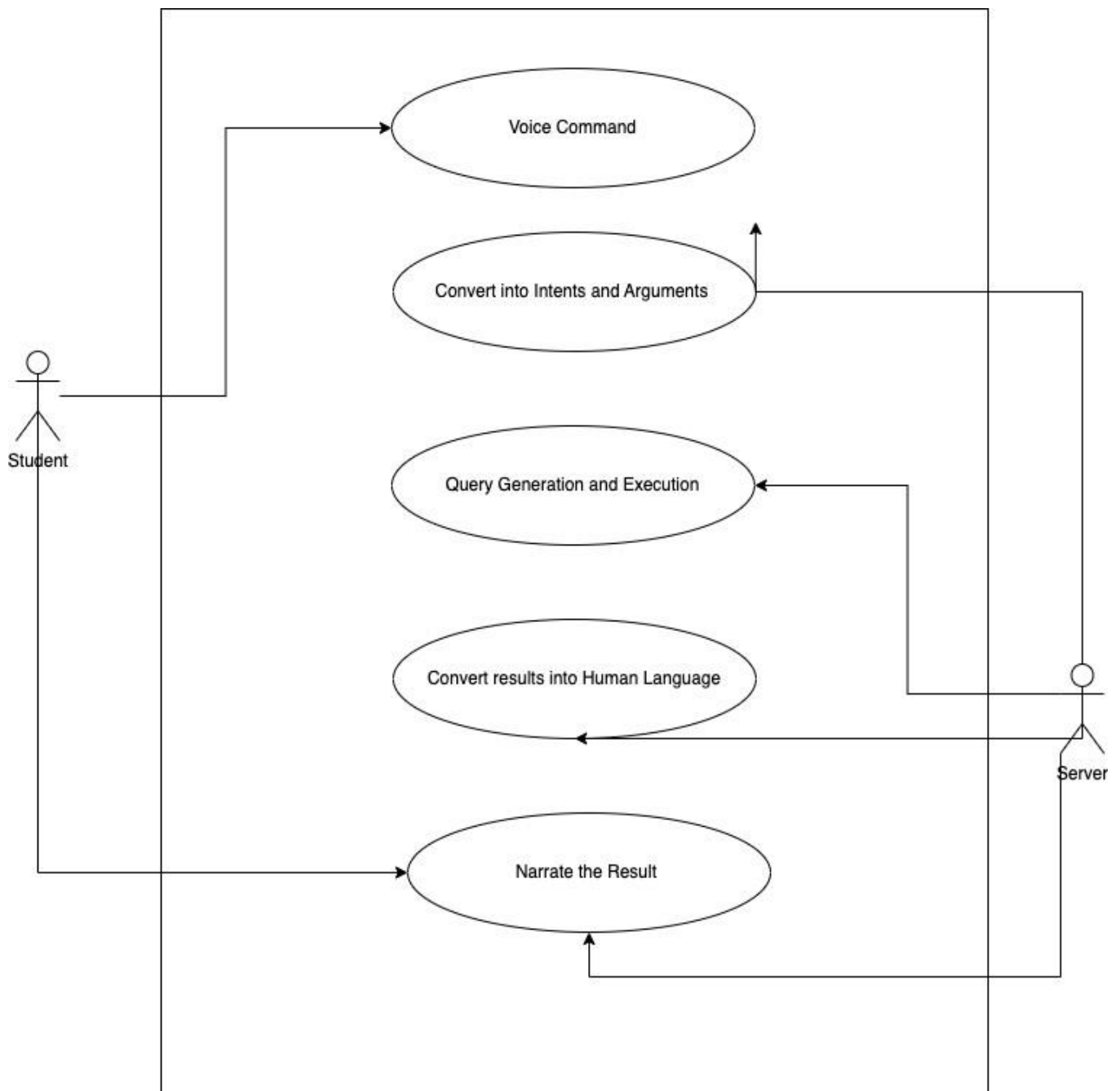
Usage Limitations

- Only applicable to Students for now. (May Extend to Professors also. Depends on work and feedback)

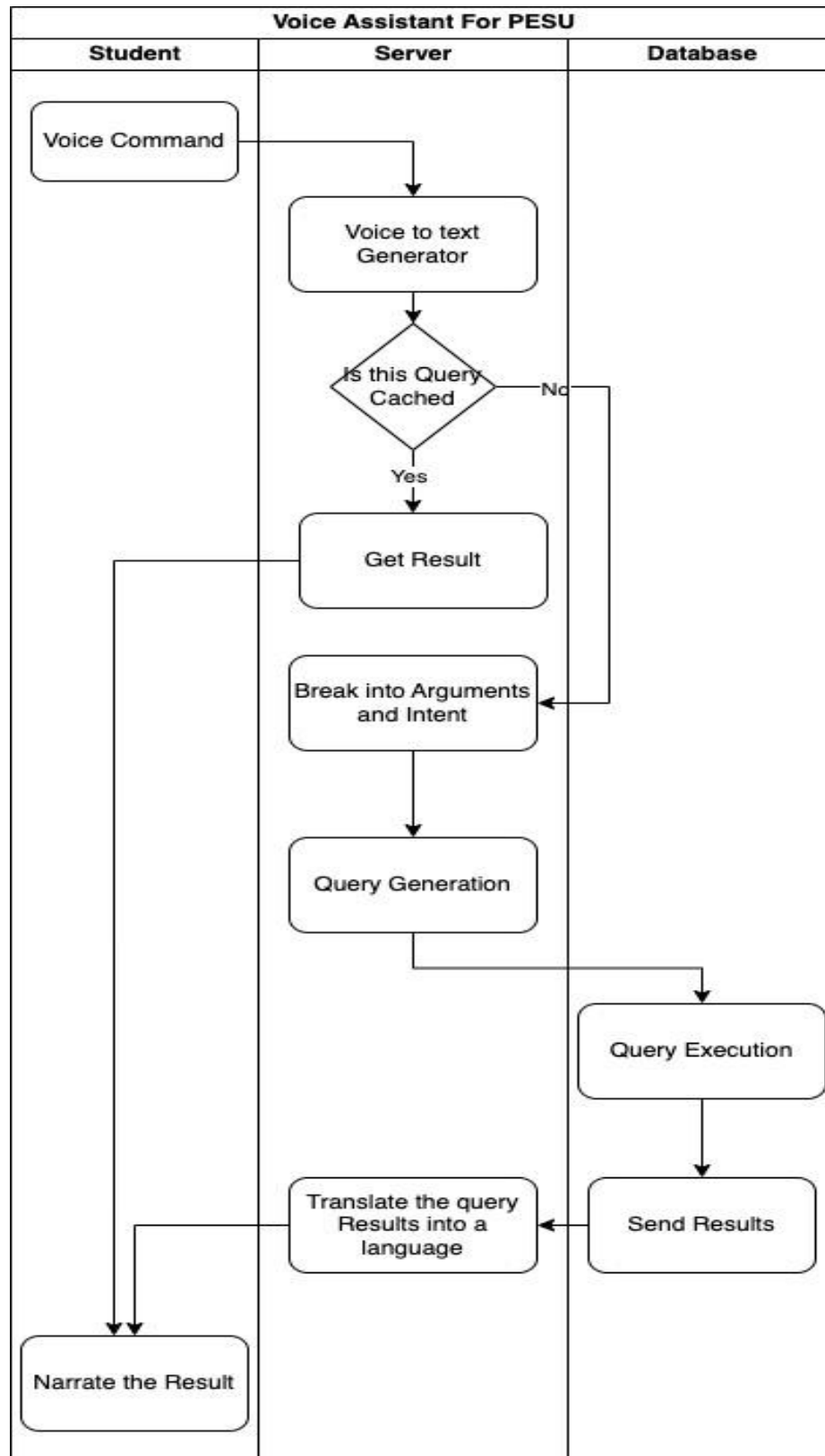
6.2 High Level System Design



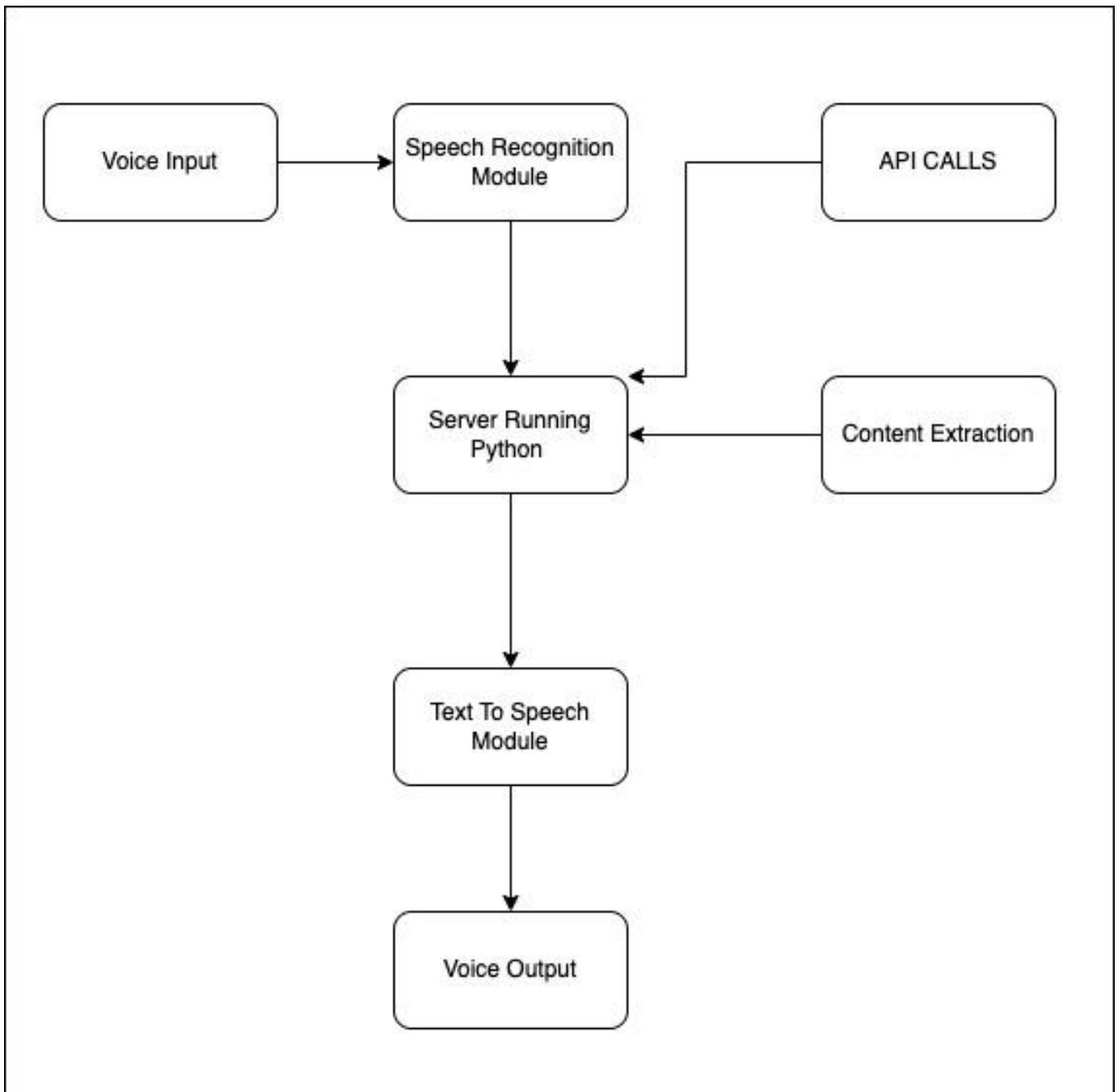
6.3.1 Use case Diagram



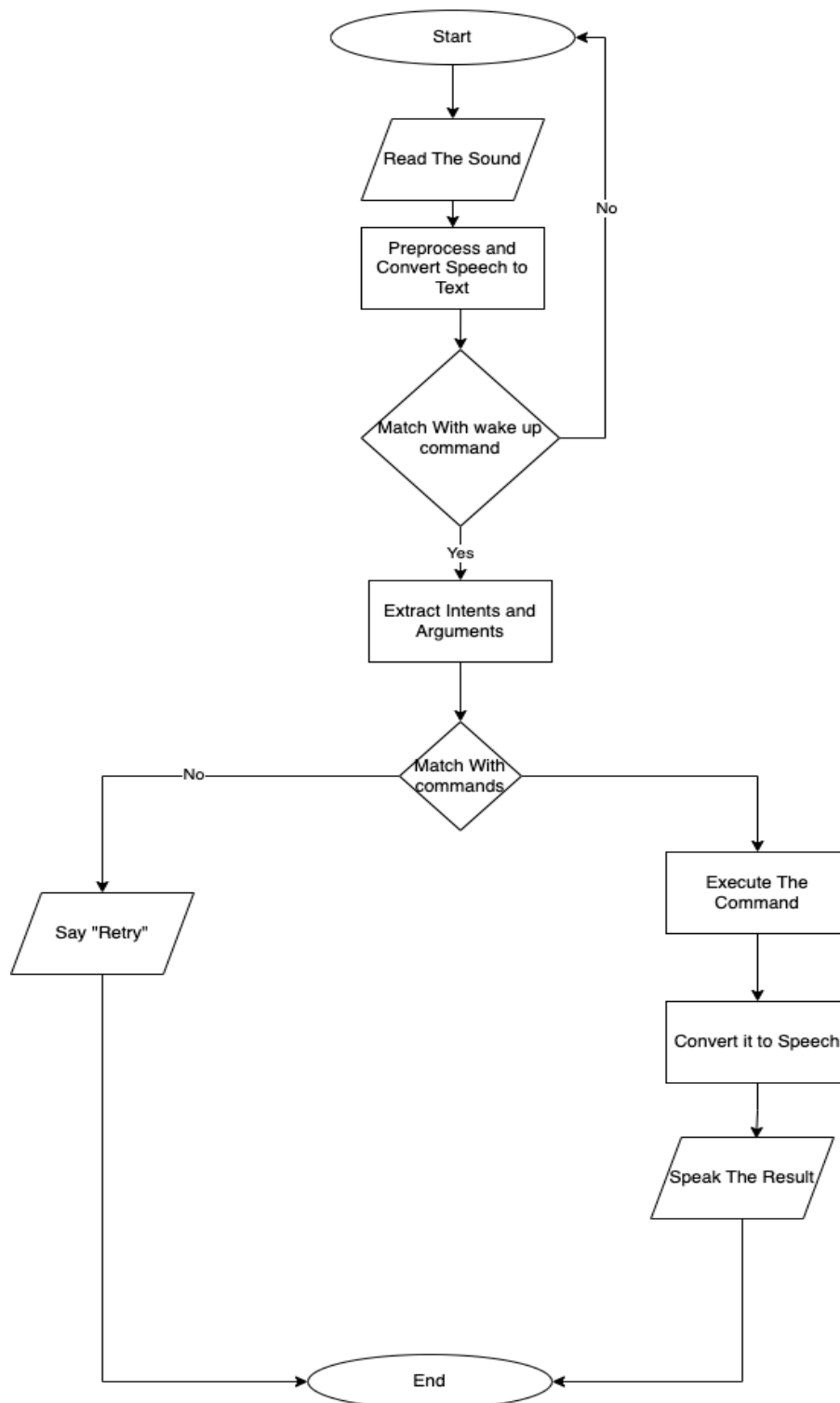
6.3.2 Swim Lane Diagram



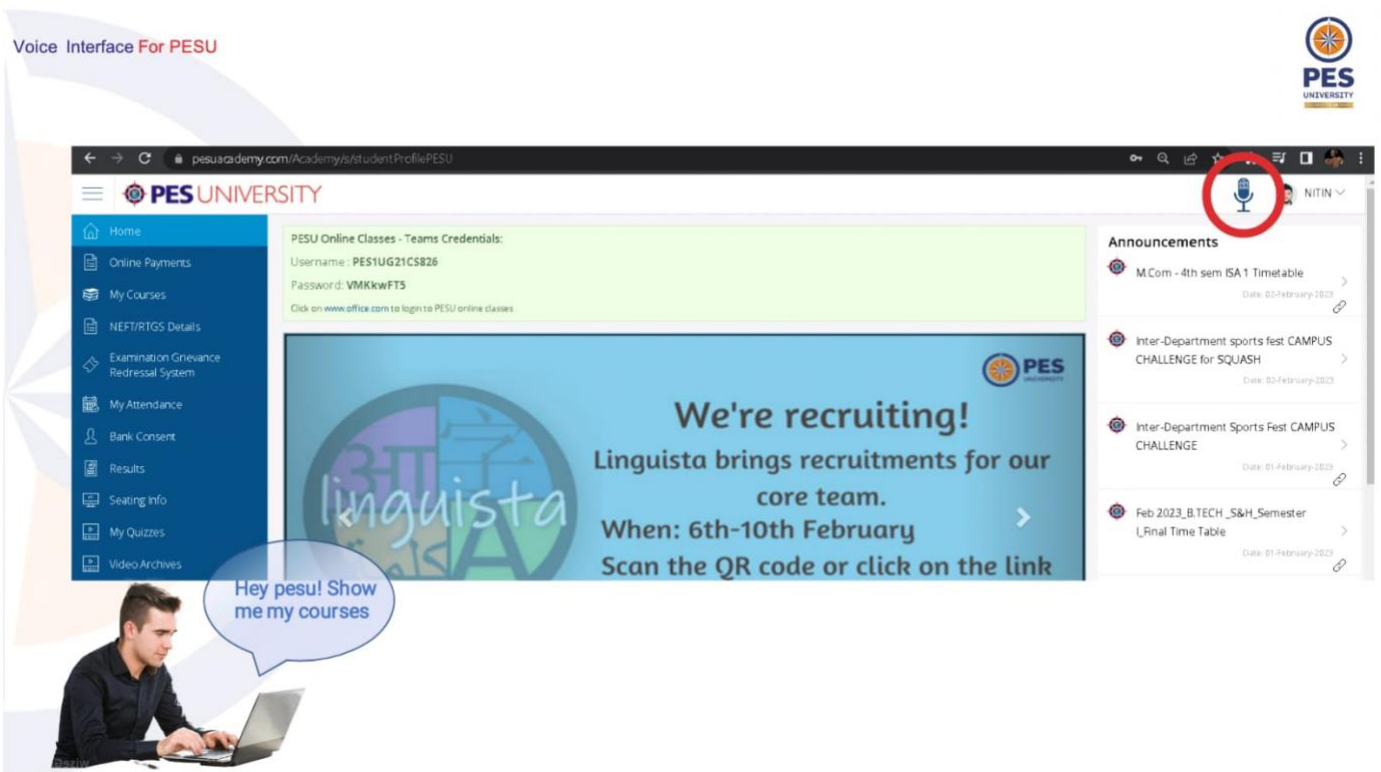
6.3.3 System design architecture for voice assistant:



6.3.4 Flow chart for voice assistant working:



6.3.5 User Interface Diagrams



CONCLUSION OF CAPSTONE PROJECT PHASE-1

7.CONCLUSION

As Discussed by panel members and our capstone faculty, we have finalized the features what to use etc.

Main Key features of our project are:

- Getting information from database
- Getting voice from user and converting it to text,
- Pre-processing it and turning it into queries

8.PLAN OF WORK FOR CAPSTONE PROJECT PHASE-2

- The main thing of part of our project is NLP, we are going to build a model for converting student query into executable query.
- We are going to design a program to pre-process all the data that we have extracted from table.
- Next part of our project is going to be database design and security, that is encryption
- We must look into handling of cached data.

REFERENCE/ BIBLIOGRAPHY

- Artificial_Intelligence-based_Voice_Assistant - Subhash S, Ullas A, Santhosh B, Siddesh S
- A Framework for System Interfacing of Voice User Interface for Personal Computers Preet Dabre, Rohit Gonsalves, Raj Chandvaniya, and Anant V. Nimkar Department of Computer Engineering Sardar Patel Institute of Technology Mumbai, India
- Table Detection and Extraction using OpenCV and Novel Optimization Methods – Nidhi, Karandeep Saluja, Asmita Mahajan, Akash Jadhav, Nakul Aggarwal
- A Method for Voice Activity Detection using K-Means Clustering - Atul Rohit Agarwal, Sourabh Tiwari, Sudhakar M S, Sankar Ganesh S
- Artificial_Intelligence-based_Voice_Assistant - Subhash S, Ullas A, Santhosh B, Siddesh S
- Table Detection and Extraction using OpenCV and Novel Optimization Methods – Nidhi, Karandeep Saluja, Asmita Mahajan, Akash Jadhav, Nakul Aggarwal