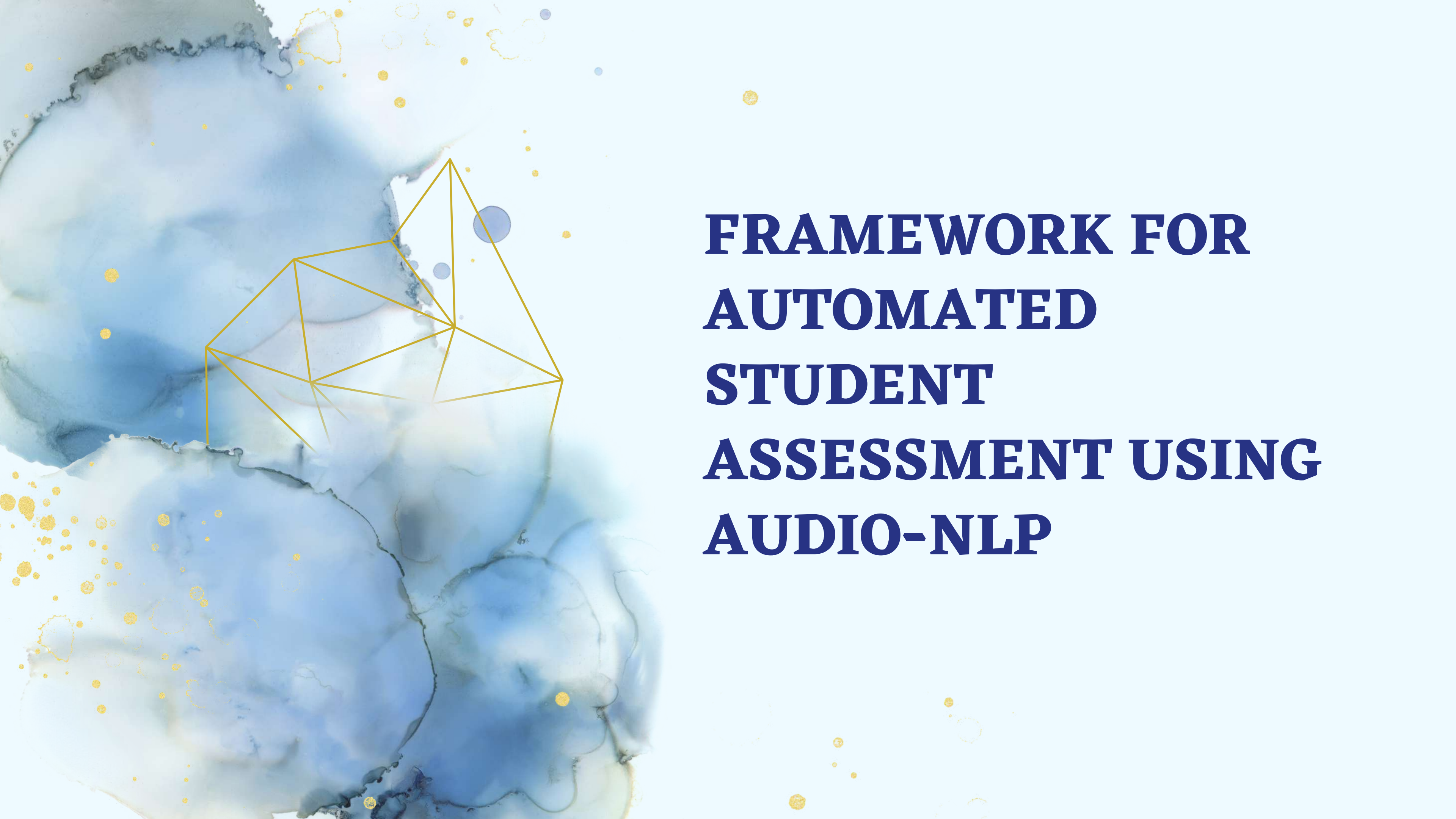




GRADE MATE

TEAM 33



FRAMEWORK FOR AUTOMATED STUDENT ASSESSMENT USING AUDIO-NLP

TEAM MEMBERS

DEVISETTI SAI ASRITH (2021111022)
SIDHI SANKALP PANDA (2021111024)
PACHAVA PAVAN (2021101013)
THOTA VENKATA SAI LAKSHMI GEETHIKA (2021101020)

Brief Problem Statement

- This application is for students to improve their skills by ongoing day-to-day assessment and being supervised by an experienced teacher. Doing this physically is time-consuming and is an administrative burden.
- The current work proposes to automate the test taking/grading process by using the audio of the student-teacher interactions during class, to perform this assessment. The work shall start with a limited set of known questions with clear answers in the form of multiple choices or True/False. Based on progress of work, exploration into use of semantic matching for allowing rephrasing of the questions by the teacher may be carried out.

To summarize the main features of this application:

- Automate the student assessment based on the student-teacher interactions during class.
- Start with basic MCQ's and True/False questions and further expand the work to long answer questions.

PROFILE OF USERS

The UI developed is meant only for demonstrative purposes only, hence has one end user with a specific use case - the development team to demo to an audience.

As such the user must be capable of using the system by providing input audio and receiving output evaluation, through a clean self-explanative UI.

Users:

- Teachers and students: Teachers and students could be the primary users of the system, using it to assess the spoken responses of the students and provide feedback on their language skills.
- Researchers: Researchers in the fields of speech and audio processing, natural language processing, and educational technology could use the system for research purposes, including evaluating the system's effectiveness, exploring new approaches to language learning and assessment, and contributing to the development of the system.
- Software developers: Software developers could use the system to build new applications and services based on the technology, or contribute to the development of the system as an open-source project.

PROJECT MODULES

User Management

- Recognize the question teacher asked during the class interaction.
- Take input audio.
- Display feedback of the candidate.

ASR Module

- Conversion of user audio into text.
- We also identify which user is speaking in the audio file.

Evaluation Module

- Find the corresponding answer of the posted question.
- Evaluate based on the scheme and produce the score.

TIMELINE HISTORY

UNDERSTOOD THE
PROJECT CONCEPT AND
DID SCOPING.

WEEK 1

IMPLEMENTATION OF
THE ASR MODEL TO
IDENTIFY TRUE/FALSE
AUDIO.

WEEK 3

DEVELOPED THE BASIC
UI.

WEEK 5

DOCUMENTATIONS AND
SELECTION OF ASR
MODULE WHICH SUITS
OUR PROJECT.

WEEK 2

DEVELOPED THE
FRAMEWORK FOR THE
UI.

WEEK 4

TIMELINE HISTORY (contd.)

CHANGED THE ASR
MODEL FROM KALDI TO
WHISPER.

WEEK 6

COMPLETE THE
FIRST STAGE
IMPLEMENTATION OF
TRUE/FALSE
QUESTIONS.

WEEK 8

INTEGRATING BACKEND
AND ASR MODEL FOR
ONE WORD QUESTIONS.

WEEK 10

IMPLEMENTING THE
ASR MODEL SO THAT
SOUNDEX CAN BE USED
ALONG WITH IT.

WEEK 7

INTEGRATING BACKEND
AND ASR MODEL FOR
TRUE/FALSE
QUESTIONS.

WEEK 9



MILESTONES

SPRINT 1

PROJECT SCOPING AND CONCEPT ✓

SPRINT 2

REQUIREMENTS COLLECTION ✓

SPRINT 3

MAKING THE ASR MODEL WORK AND BUILD UI ✓

SPRINT 4

BACKEND FOR WEB APPLICATION ✓

SPRINT 5

DEVELOPMENT OF DASHBOARDS ✓

SPRINT 6

TESTING THE INTEGRITY OF ASR MODEL AND DATABASE ✓

SPRINT 7

DEPLOYMENT ✓

Feature Highlights

The developed System contains the following features

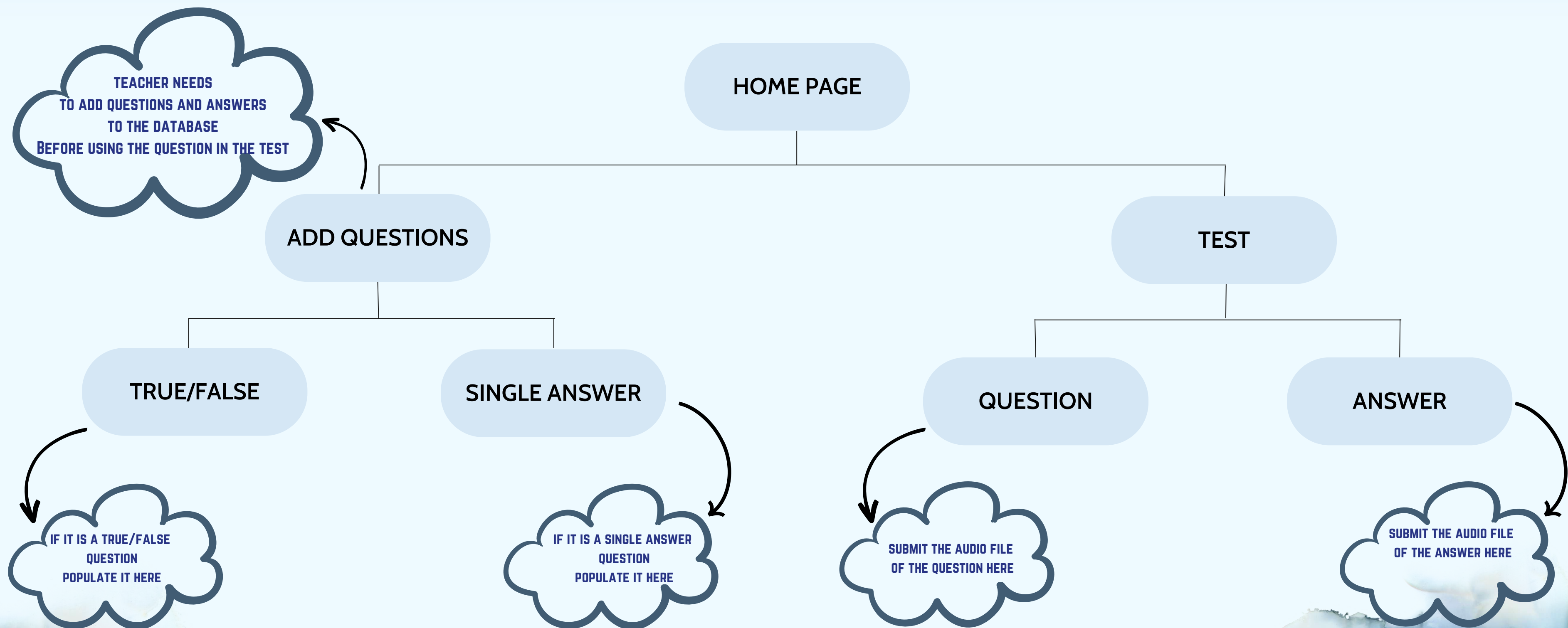
FRONTEND/UI

- Enter input audio.
- Perform API calls.
- Display text generated and output evaluation.

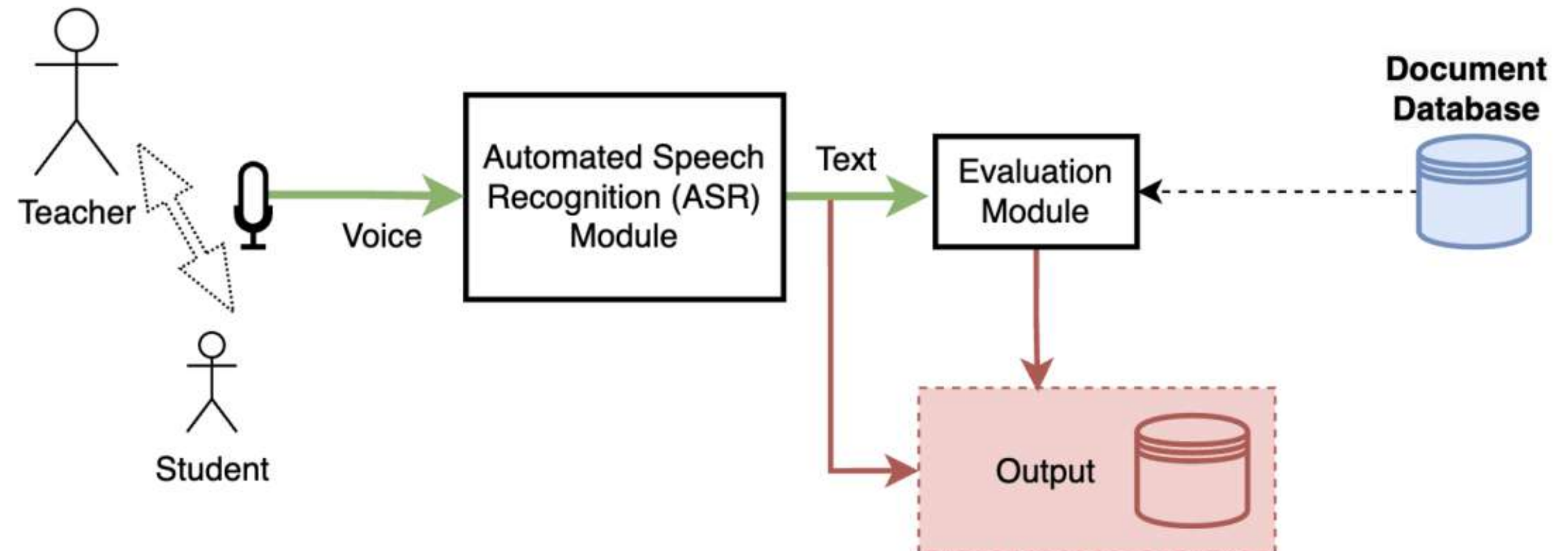
BACKEND/API

- Receive input audio.
- Run ML inference models to convert audio to text, check if text matches with curated set of questions and answers, and finally generated output evaluation.
- Send back output evaluation and converted text.

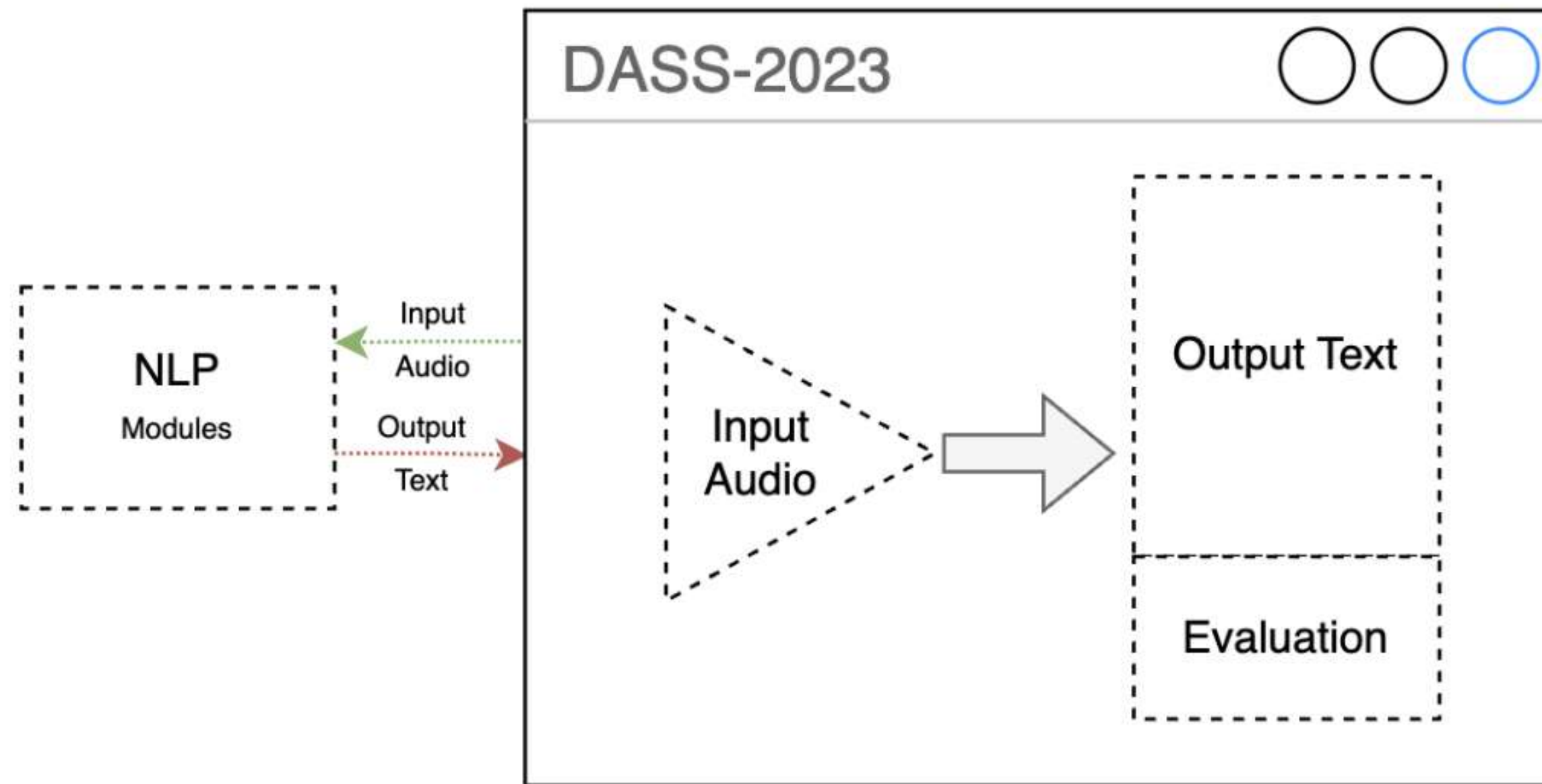
Hierarchy of UI



OVERALL SYSTEM ARCHITECTURE



UI/BACKEND PROCESS FLOW



Development tools

ASR Models

Whisper
(along with Soundex)

Conversion of
speech-to-text.

Python

Python 3.9.12

Data base
connectivity,
gluing together
different APIs.

Word2Vec

word2vec 0.11.1

Generates word
embeddings and maps
each word to a vector
which helps in determining
closeness of sentences .

Flask

Flask 1.1.2

Quickly create web
APIs that implement
best practices by
default.

React

React 18.2.0

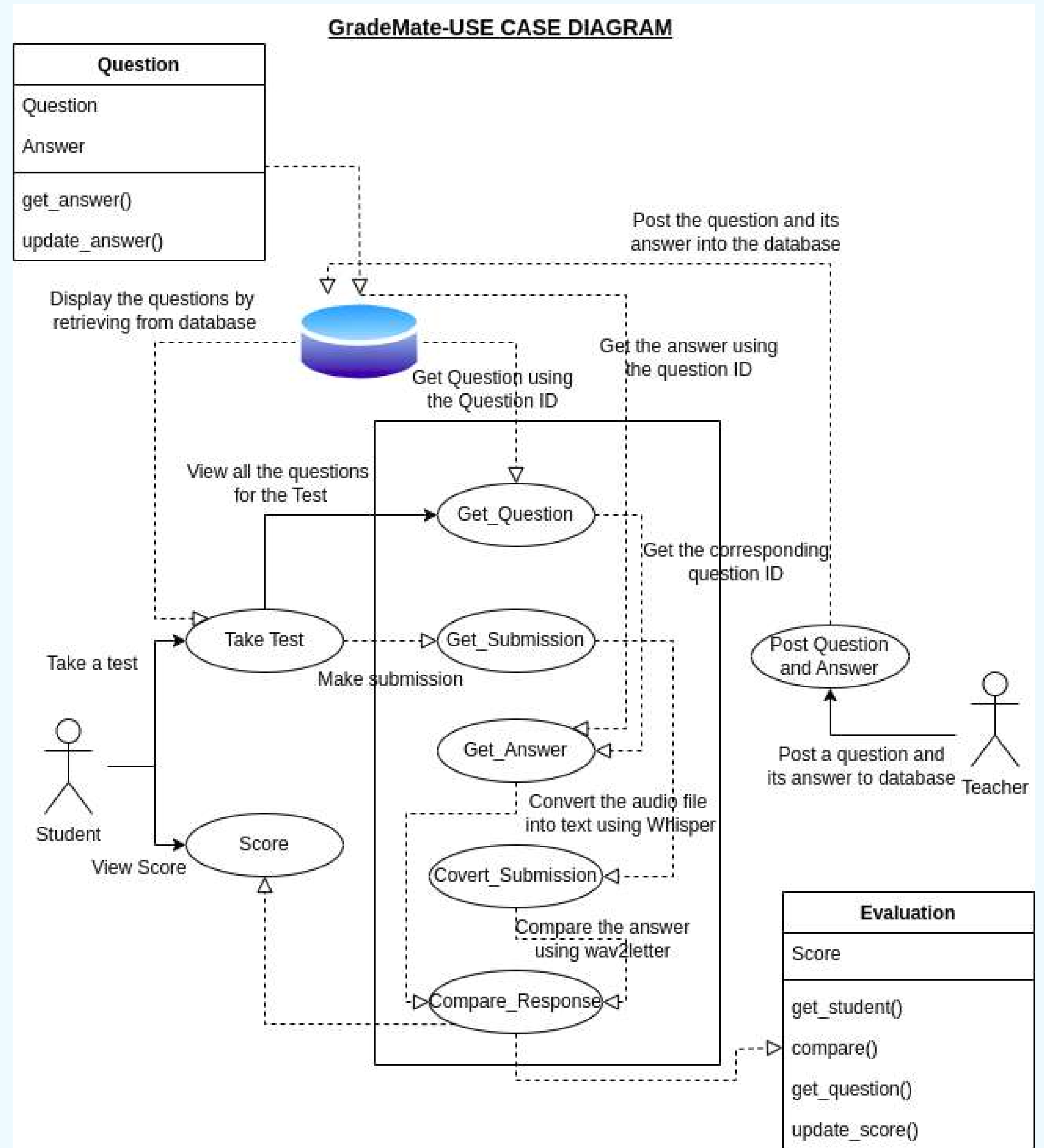
Creates a user-
friendly interface.

MongoDB

MongoDB 6.0

Store data
regarding the
correct answers.

UML USE CASE DIAGRAM



ASR MODELS

What are ASR Models?

Automatic speech recognition or speech-to-text models aim to convert a raw audio file into a sequence of corresponding words. For the current model under deployment, we need to check the required specifications that suit our project.

The best available ASR Models are;

- **Mozilla Deep Speech**
- **Kaldi**
- **Wav2letter++**
- **TensorFlow ASR**
- **Open Seq2Seq**

SELECTION OF ASR MODELS

MOZILLA DEEP SPEECH

- This model is trained using machine learning techniques.
- It is implemented using TensorFlow.
- We can also easily integrate it into your other machine learning projects that we are having on TensorFlow.
- Allows us to use a pre-trained model or can use train our own model.

KALDI

- Kaldi is mainly coded in C++ wrapped with bash and python.
- The best feature is the Extensible design.
- For the basic usage of converting speech to text, then we'll find it easy to accomplish that via either Python or Bash. We also have Kaldi Active Grammar, which is a Python pre-built engine with English-trained models already ready for usage.

Wav2letter++

- This model is completely built on C++ and fastest ever.
- The machine learning library used is Flashlight.
- Allows us to use a pre-trained model or can use train our own model.

TensorFlow ASR

- Very recently developed model.
- More up-to-date than DeepSpeech..

Open Seq2Seq

- It is developed for sequence-to-sequence model training.
- Even in this model, we can train a model or use a pre-shipped model.
- It is more up-to-date than Mozilla's Deep Speech, Wav2letter++..

MODEL	COMMENT
Deep Speech	This is one of the oldest models whose up-gradation is almost stopped. So in case of any error, it may cause difficulties.
Kaldi	One of the most commonly used ASR models. Can be considered an ASR model for our project.
Wav2letter++	Can be considered an ASR model for our project
TensorFlow ASR	Still developing ASR model, even though better than others.
Open Seq2Seq	Better than many, but its updates were stopped and for out basic need we don't need seq2seq model

Out of all, we tried to go with Kaldi.
But, we faced some challenges with it. So, we selected "Whisper"



Whisper Module

- Whisper is an automatic speech recognition (ASR) system trained on multilingual and multitask supervised data collected from the web.
- We show that the use of such a large and diverse data-set leads to improved robustness to accents, background noise and technical language.
- Moreover, it enables transcription in multiple languages, as well as translation from those languages into English.

word2vec

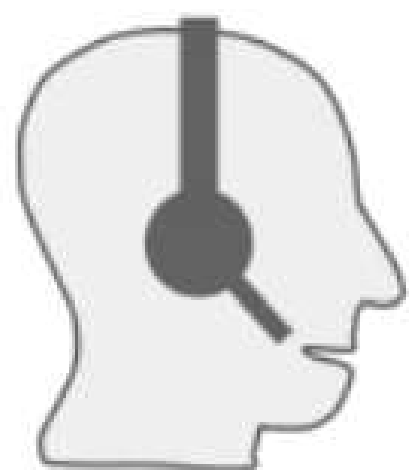
- Word2vec is a popular natural language processing (NLP) technique used for generating word embeddings. It is a neural network-based model that learns to represent words as vectors in a high-dimensional space, such that words with similar meanings are closer together in that space.
- The word2vec model can be trained using large amounts of unlabeled text data, such as Wikipedia articles or news articles. The model is trained to predict the probability of a word given its context (the words that appear around it in a sentence). This is known as the skip-gram model.
- Once the model is trained, each word is represented as a dense vector in the high-dimensional space, with each dimension representing a feature of the word. Words with similar meanings are located near each other in this space, allowing for easy computation of semantic similarity between words.
- The word2vec model has many applications, such as in information retrieval, sentiment analysis, and machine translation. It has become a popular tool in the field of NLP due to its ability to capture the semantic relationships between words and its effectiveness in improving the performance of many NLP tasks.



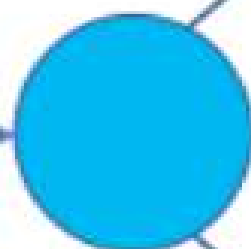
Speech



Text



Speech input



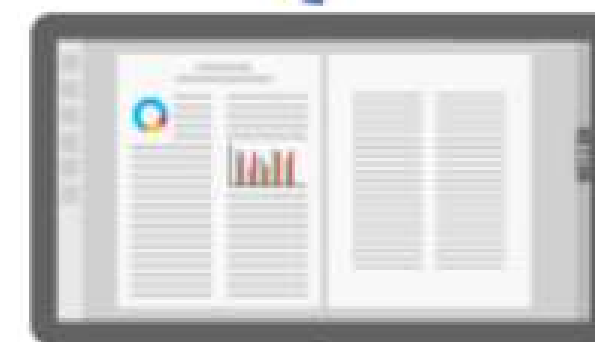
Whisper / normal voice classifier

Whisper voice recognizer

Normal Speech Recognizer

Newline, delete, menu ..

commands



Text input

A quick brown fox ...



“

The motive behind our project

The motive behind the project "AudioNLP for students assessment" could be to explore the potential of using speech and audio processing techniques, including natural language processing (NLP) and automatic speech recognition (ASR), to develop a system for assessing students' performance based on their spoken responses.

This project could have various objectives, including improving the accuracy and efficiency of student assessments, reducing the workload of human assessors, and providing more objective and standardized evaluations of students' spoken responses.

The use of NLP and ASR technologies in this project could enable the system to analyze and interpret the spoken responses of students, including their grammar, vocabulary, pronunciation, and fluency. This could provide valuable insights into the strengths and weaknesses of students' language skills, and help identify areas for improvement.



GRADE MATE

WEB APPLICATION

A Friendly Tool for Managing and Grading Student Assessments!

Add Questions

[Click here to populate the questions database](#)

Take Test

[Click here to view the demo](#)

Enter the Question and Answer

Question

Is this a screenshot?

Answer

Choose your options:

☒ True

☐ False

Post

Cancel add request:

Back

OK

Enter the Question and Answer

Question

Is this a screenshot?

Answer

Choose your options:

☒ True

☐ False

Post

Cancel add request:

Back

A friendly tool for managing and grading student assessments.

Add True/False
Question

Add One Word
Answer Question

localhost:3000 says
Posted Sucessfully

OK

Enter the Question and Answer

Question

What is this for?

Answer

Project

Post

Cancel add request:

Back

Welcome to the test, let's start!

Question:



Answer:



Evaluate

▶ 0:00



▶ 0:00



Question:

Am I okay?

Answer:

False

Sumbitted Answer:

forit

Result:

Correct

Welcome to the test, let's start!

Question:



Answer:



Evaluate

▶ 0:00

🔊

▶ 0:00

🔊

Question:

Is this working?

Answer:

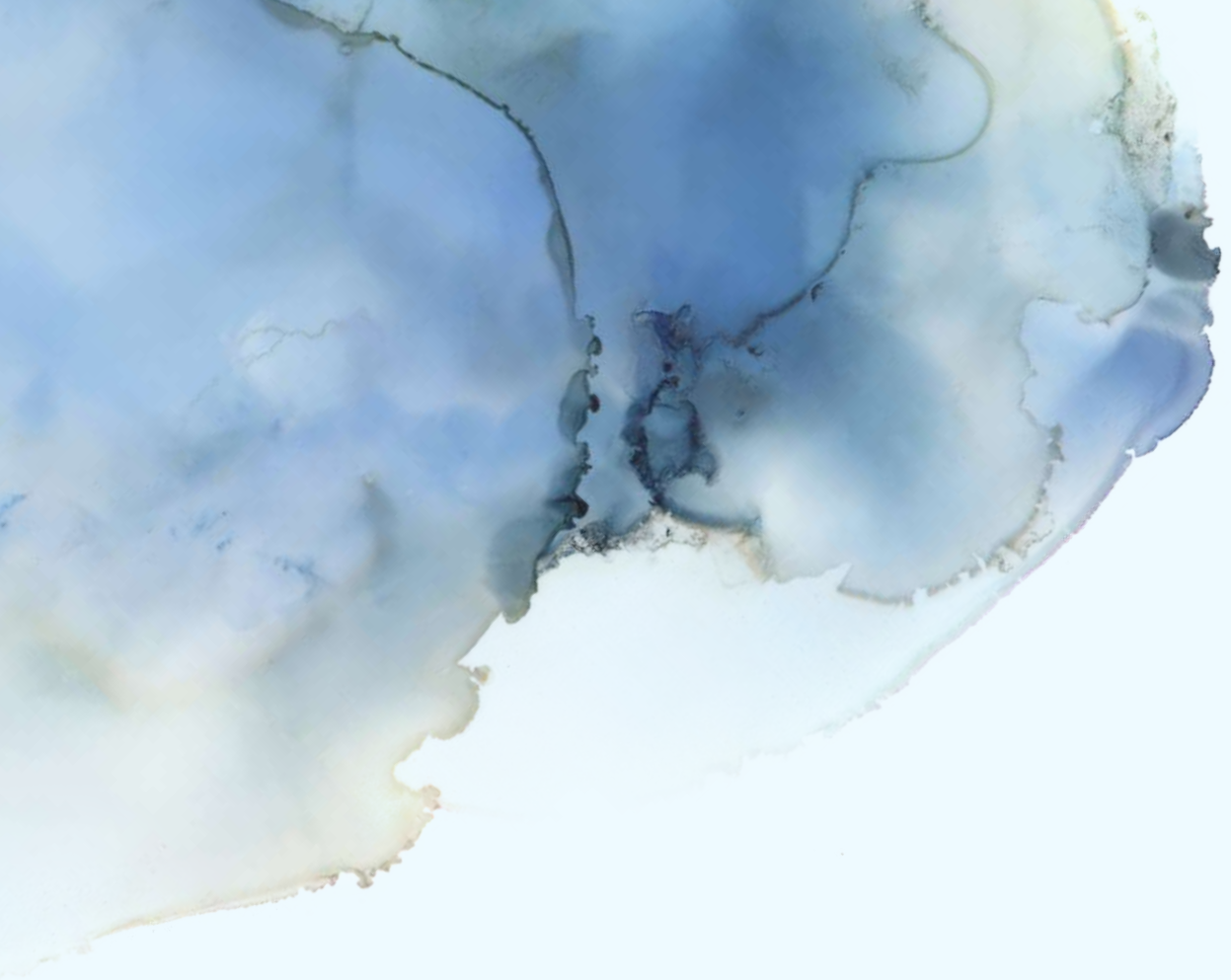
Probably

Sumbitted Answer:

Probably

Result:

Correct



ALSO, HERE IS THE LINK FOR OUR DEMO PRESENTATION

https://iiitaphyd-my.sharepoint.com/:v:/g/personal/pavan_pachava_students_iiit_ac_in/ESsRq0lBOt9Ilj2bYFidoUQBswiOUQcq6fLo1jwm2Hvc9Q?e=7HJl20



Roles

01

**Sidhi Sankalp
Panda**

Developer
Tester

02

**Geethika
Thota**

Developer
Tester

03

**Sai Asrith
Devisetti**

Developer
Tester

04

**Pavan
Pachava**

Developer
Tester



THANK YOU

TEAM 33