Developing a Virtual Reality System Integrated with Large Language Models for Real-Time Evaluation and Feedback to Improve Public Speaking Skills

Yuantu Chen [1], Justin Dang [2]

[1] The Bush School, 3400 East Harrison St, Seattle, WA 98112
[2] California State Polytechnic University, Pomona, CA, 91768

ABSTRACT

Public speaking skills are considered both difficult as well as anxiety inducing for many, which in our society, dominated by frequent communication and presentation, can be problematic as well as prohibitive[1]. We propose a system using VR technology and AI large language models in order to help assist users in practicing their public speaking skills [2][3]. Users will be situated in a virtual environment, and an AI model will grade their speech and provide them notes on improvement. This required overcoming several design challenges such as prompt engineering our LLM as well as speech transcription. We performed an experiment in order to test our LLM model by having it grade varying types of speeches. Analysis of the data supports the idea that our model is consistently evaluating user speeches at the quality that we expect it should, although there are some improvements we could make to the AI model to improve its evaluation quality even further.

KEYWORDS

Public Speaking, Virtual Reality, AI Feedback, Speech Evaluation

1. INTRODUCTION

Public speaking is one of the most significant and widely occurring fears present in our society, from all levels. According to several studies, roughly 75% of the world's population have some level of glossophobia and have anxiety when it comes to speaking in public. This fear of public speaking is significant in a world that is driven largely by presentation and communication. It is estimated that those with a fear of public speaking receive 10% less wages and are passed on for promotions by 15%. According to surveys, 45% of individuals interviewed believed that their fear of public speaking negatively impacted their ability to advance their career. Most importantly, a reported 60% of people interviewed expressed that they wanted to improve their public speaking skills. Public speaking is pervasive in our society, and is utilized by students all the way to those brought on to speak in governments or organizations such as the United Nations. It is important that we as a society try to solve this widespread problem related to fear of public speaking. Enhanced communication skills would increase workplace and academic productivity immensely, positively impact career prospects for many individuals with a fear of public speaking, and alleviate generalized anxiety in many peoples, thereby supporting many's mental health.

The paper AI-Powered Analysis with Facial Expression Recognition by Padia et al. focuses on facial expression recognition and analysis technology in order to evaluate users' public speaking skills [4]. We believe that our approach to evaluation is more comfortable for the user and more applicable to consumers. Secondly, the paper Virtual Reality Public Speaking Training by Bachmann et al. focuses on utilizing VR environments, much like our approach, to improve public speaking skills. However, their approach does not incorporate AI functionality to improve the user experience; as such our approach may be considered an extension of their work. Lastly the paper Multimodal Visual Analytics System by Huang et al. utilizes mainly visual analysis techniques in order to evaluate users' speeches [5]. However, we believe that the lack of a virtual environment in their case allows for users to more properly immerse themselves in a public speaking scenario and give more natural performances.

We have devised an application that utilizes both virtual reality technology and generative AI systems in order to help people practice their public speaking skills in an isolated and controlled environment. The main goal is for us to successfully recreate a high-pressure, immersive environment in a virtual reality environment. Virtual reality was chosen as it allows users to simulate environments that they typically would not have access to. In our case, we can easily simulate a crowded environment that the user can practice in without bothering other people. A user in our program would be in an environment similar to a UN speaking podium and can practice their speech in isolation. The system will record their speech through their microphone and transcribe it. A large learning model (LLM) will then grade their response based on the structure of their speech based on several criteria and provide the user a final score at the end[6]. We believe that this virtual public speaking simulation helps users overcome their public speaking anxiety through repeated, controlled exposure. Utilization of AI, specifically large language models, allow us to go past traditionally slow and somewhat unreliable systems of natural language processing such as syntax analysis.

For our experiment, we wanted to test how flexible and how accurate the AI grader was in determining the quality of users' speeches. To do so, we tested the model on a wide variety of speech formats that vary in length, quality, word choice, and other metrics. Our goal was to predict the grade or at least establish what a human grader would reasonably assign to each speech, and determine how the AI model graded in comparison. We deemed an AI assigned grade that was close to our expected grade to be considered an accurate assessment. In 3 out of 4 test cases, the AI grades matched our expectations. However, the AI only gave broad grades and did not prefer to use plus or minus marks to give minute evaluations. Although mostly accurate, we believe some more fine tuning is needed in order for the model to provide more holistic grades that better reflect the user's performance.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Accurate AI Grading for Public Speaking

Special emphasis is placed on our AI-based analysis and grading system. This is because we want to correctly identify shortcomings in the user's speaking style, and not fall into typical problems with utilizing LLMs such as hallucinations, overconfidence, or overly positive outputs. Utilizing prompt engineering techniques, we were able to essentially program LLM models such as ChatGPT to assume a different mode of thinking and output style when taking into account the

user's speech. After rigorous testing, we believe that the model is capable of consistently critical and proper analysis.

2.2. Designing VR Environments for Effective Speech Practice

Another concern that some may have regarding the effectiveness of our model is whether or not the environment simulated in the virtual reality headset is sufficient enough for users to properly give speeches in. We believe that the usage of a virtual United Nations stage correctly encapsulates an environment in which individuals with or without public speaking anxiety are forced to give their best public speaking performance and practice. However, it is important that this environment is not too demanding on those with a fear of public speaking such that they do not want to participate in it. We believe the use of a virtual reality headset and the fact that the application is strictly one-user only allows the user to practice on their own terms.

2.3. Ensuring Accurate Transcriptions for AI Speech Evaluation

An AI model will grade the user's public speaking ability based on a transcript of their speech as done through the app. One concern is whether or not the transcript is correctly generated. An improperly generated transcript will affect the grade given, even if the speech itself could be otherwise considered to be of high quality. Inversely, a poorly worded speech may be falsely flagged as well done. We utilize Whisper for our transcription technology, which is a popular and consistently accurate transcription library offered by AI technology company OpenAI. As such, we believe our transcriptions are sufficiently accurate.

3. SOLUTION

The user is met with a podium with a display that includes the Start recording buttons which the user can press to begin recording their speech. If the user does not possess a VR headset, they can use the w, a, s, and d keys to move around the theater scene to simulate giving a speech in front of a realistic audience. Once the user has finished giving their speech or argument, they can maneuver toward the podium and press the Done button, ending the recording session. In the background, the Unity audio file is converted to a .wav file and saved to the Recordings folder. The program will then do another conversion, this time from an audio file to a text-based string which ChatGPT can then review [7]. Finally the ChatGPT response to the user input will be printed out on the display located on the podium. For the user, after a quick second the podium display will light up with thorough feedback as well as a letter grade. There is a scroll bar on the right hand side which the user can drag with their cursor to view the full response. There is no extra screen in this prototype. To begin practicing another speech, the user will have to resort to restarting the program.
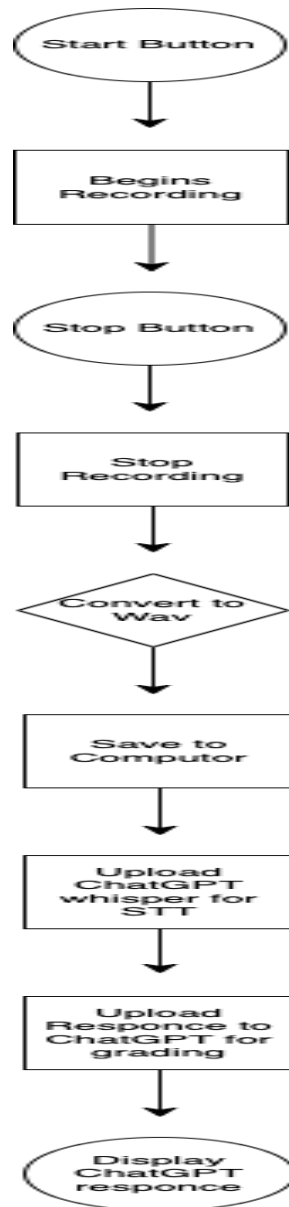
Figure 1. Overview of the solution

The first component is converting a Unity audio clip into a .wav file. The purpose of the conversion is that ChatGPT's whisper module does not accept recordings in the form of Unity audio clips. We did not introduce any backend or database services. However, we manually converted Audio Clips to a .wav file using the code we wrote.
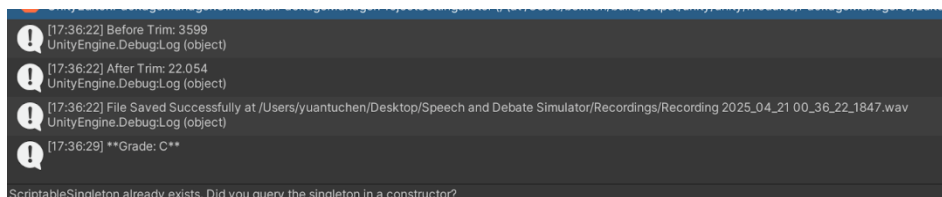

Figure 2. Screenshot of the first component

```
// Creates a folder at head of project and saves AudioClip as .wav, returns the file path
public static string SaveRecording(AudioClip audioClip)
{
    if (!audioClip)
    {
        Debug.Log("Error with given AudioClip in AudioManager.SaveRecording().");
        return "";
    }

    DirectoryInfo directoryInfo;
    // Checks if folderName exists at head of this project folder, the one that contains the asset folder
    if (!Directory.Exists(folderName))
    {
        // CreateDirectory will create folders at the head of this project folder, the one that contains the asset folder
        directoryInfo = Directory.CreateDirectory(folderName);
    }
    else
    {
        directoryInfo = new DirectoryInfo(Directory.GetCurrentDirectory() + "/" + folderName);
    }

    string filePath = Path.Combine(directoryInfo.ToString(), fileName + " " + DateTime.UtcNow.ToString("yyyy_MM_dd HH_mm_ss_ffff") + ".wav");

    // Delete the file if it exists.
    if (File.Exists(filePath))
    {
        File.Delete(filePath);
    }

    try
    {
        WriteWAVFile(audioClip, filePath);
        Debug.Log("File Saved Successfully at " + filePath);
        return filePath;
    }
    catch (DirectoryNotFoundException)
    {
        Debug.LogError("Persistent Data Path not found!");
        return "";
    }
}
```

Figure 3. Screenshot of code 1

This function saves a Unity AudioFile to our project as a .wav. Our SaveRecording function includes AudioClip and audioClip as parameters. The first if statement checks for the existence of the trimmed audio clip that was recorded from Unity, if it does not, an error message will be printed. The code checks whether the recording folder exists or not; if not, it will create a recording folder for storing recordings. We then give a unique name to the recording file with the date and time. Next, we convert the Unity audio clip into a .wav file and save it at the filepath.

ChatGPT whisper allows us to convert our recorded audio into text, which ChatGPT can then grade. ChatGPT can only grade text-based files, not audio. We used a api created by OkGoDoIt that allows us to use ChatGPT in C#. It handles all the requests we make to ChatGPT. This component uses Natural Language Processing which is a field in AI that allows computers to interpret audio bytes/information into a string of text [8].

```
public async void GPTGradeAudio(string filePath, float clipLength)
{
    string resultText = await api.Transcriptions.GetTextAsync(filePath);

    GPTGradeScript(resultText, clipLength);
}
```

Figure 4. Screenshot of code 2

We create another function with filePath and clipLength as our two parameters. FilePath is where ChatGPT whisper gets the location for the converted audio file. ClipLength is requested in the GPTGradeAudio function so that we could pass it onto GPTGradeScript. Inside our function we call another function, api.Transcriptions.GetTextAsync. This function is responsible for converting our audio file into a string or text format. The function has multiple parameters such as language, prompt, temperature. Language is used to define the language of the speech in the audio file. Temperature allows us to define the consistency of the transcription. Prompt is used to provide context for the audio clip we are sending to OpenAI's Whisper API [9]. Finally we call the previous function responsible for grading the user input.

The third component we used was the ChatAPI that ChatGPT provides. it allows us to provide a prompt for ChatGPT to follow(eg. You are a judge of a speech and debate competition) which it will use to respond to any question/message we send to ChatGPT. We use OkGoDoIt's API to access the ChatAPI. This component relies on Natural Language Processing which is a concept that allows computers to understand our language.
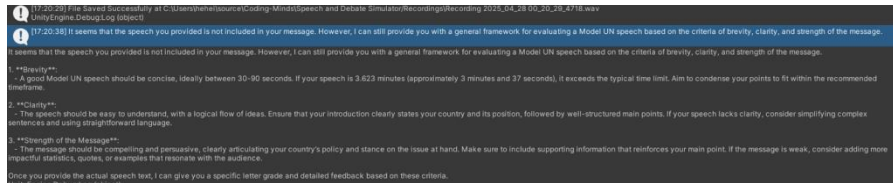

Figure 5. Screenshot of the thrid component

Figure 6. Screenshot of code 3

The function takes in a string and a float; it uses these parameters to create a message with information for ChatGPT to use for grading. The string contains the transcription of the audio file that we received from the second component of our project. The first line of the function creates a chat containing the prompt ChatGPT will use and the message that ChatGPT will respond to. Here we can also specify a Chat model we want ChatGPT to use as well as a temperature [10]. We use temperature to control the consistency of ChatGPT responses. A lower temperature means more consistent answers. We then take ChatGPT's response and insert it into the display panel and debug the message in the console to see it.

4. EXPERIMENT

4.1. Experiment 1

We want to test out the accuracy of the feedback ChatGPT provides after grading a speech. It is important to conduct this experiment to ensure that our speech and debate simulator can give consistent feedback depending on the quality of the input. The feedback factor is crucial to the usefulness of program as it is what users seek after pouring their hearts out for the sake of learning and improving on their public speaking skills.

Our inputs must be varied in terms of quality and classifications. Ranging from simple everyday comments to long info dense speeches, our users will provide a great variety of inputs and our program must be able to identify whether they pertain to speech and debate. By providing a good selection of recordings, some good and some bad, we can locate areas of the program that require adjusting and make the program more robust. For the program to be most adaptable, it must be able to provide feedback for even cases that are unexpected. With an adaptable and consistent

program not only can we ensure that all users are accessible but can also open the doors for more. Varied inputs will push the program to its limits and will only allow us to push it further!

	Input	Expected Output	Actual Output
1	"The weather is pretty bad today. It won't be a big issue if I leave the blankets to dry outside though." (21 words / 7s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT recognizes this as an invalid speech. Overall: F Brevity: F Clarity: F Strength: F	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT does not consider this a valid speech. Overall: F Brevity: F Clarity: F Strength: F
2	"I had a great day! Tom said that he forgot to water the flowers, so they will probably be dead by Monday." (22 words / 7s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT does not consider this a valid speech. Overall: F Brevity: F Clarity: F Strength: F	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT does not consider this a valid speech. Overall: F Brevity: F Clarity: F Strength: F
3	"Does anyone even read these posts? I kind of just say whatever here and no one responds." (5 words / 6s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT does not consider this a valid speech. Overall: F Brevity: F Clarity: F Strength: F	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. ChatGPT does not consider this a valid speech. Overall: F Brevity: F Clarity: F Strength: F
4	"Honorable chair distinguished delegates, The first step to stopping starvation is improving access to clean water, critical for food growth. Half of the 800 million people without access to clean water live in Sub-Saharan Africa. The Kingdom of Norway believes that the WFP's efforts should focus on community-oriented aid to promote self-sufficient living. Norway thinks self-operated water harvesting devices are a good way to do this. The University of Akron in Ohio developed a water harvester that produces 10 gallons of drinking water per hour from thin air. The UN should purchase, and distribute, water harvesters to the countries most in need such as Niger, Burundi and Mozambique. For the UN to ensure long term success, the WFP should transfer harvesters and not funds to the countries in need." (129words / 43s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: B Clarity: A- Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: A Strength: B
5	"Honorable chair distinguished delegates, The United Kingdom is strongly against making visa access easier. While the death of the 39 Chinese found inside a refrigerated lorry from Bulgaria on October 23rd, 2019 is regrettable, our only option is informing potential migrants of the danger to themselves. Illegal migration into Britain is around 650,000—give or take a couple hundred thousand. Many of these enter countries on tourist visas and then stay. Changing the laws will only give smugglers more opportunities. Britain proposes the UN create translated online platforms to apply for legal visas, while also showing the dangers of illegal migration. The UN should invest in the proliferation, so this information reaches the right people. The smugglers who get past our x-ray machines, canine units, heartbeat monitors and carbon-dioxide sniffers are extremely resourceful. If we loosen visa laws, they will adapt and continue to take money from the poor but now with the white hats, we handed to them." (159 words / 43s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B+ Brevity: B Clarity: B+ Strength: B+	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
6	"We, the Arab Republic of Egypt believe that we should help rebuild the Middle East and specifically Syria. While the 580,000 casualties of the civil war in Syria are extremely regrettable, the real victims are the people trying to stay alive in the conflict zone that is modern Syria. More than 6.2 million people are displaced. 13.1 million are still in need of humanitarian assistance. The estimated unemployment rate stands at 54 percent. Also, 83.5 percent of the 19,454,263 Syrians live below the poverty line. Cities have been engulfed in crime, police stations closed down and the overall police personal dropped from 100,000 officers to 20,000. Rates of theft increased, with criminals looting houses and stores. To fix this problem we need to rebuild the Middle East by rebuilding Syria!" (139 words / 41s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: C+ Brevity: C+ Clarity: B- Strength: C+	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
7	"The Republic of Ireland declares that we need to rebuild the Middle East! Ireland believes that we should take action against the war and apply sanctions against the rebel terrorists in Syria and even resort to military action if necessary to stop the war. Syria is in such a poor condition because of the west's irresponsible behavior regarding the war. The millions who died and fled are directly our fault. This is the same level of neglect seen after the Vietnam war in 1975, when we left over 2 million as the casualty count and fled to lick our wounds and rebuild ourselves. Once the war is over, we should send financial aid to rebuild Syria again and prevent any future wars from happening." (129 words / 38s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A Brevity: A- Clarity: A Strength: B+	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
8	"Peru believes that every Arctic country has the sovereignty to use their resources. However, because of the climate changes, we should create a special committee that will discuss this subject and find ways to decrease the ecological damage. The Arctic region contains major reserves of uranium, copper, tungsten, gold, diamonds and most importantly gas and oil. In addition, it's one of the largest freshwater reservoirs in the world. Climate changes and release of contaminants in the Arctic have potential to affect European and global weather patterns. The Arctic is particularly sensitive to the effects of global warming, and icebergs are melting at a rapid rate. Scientists fear that it will cause a significant rise in sea levels around the world, and that in the 20th of the 21 century there will be no ice zone at all during the summer."(140 words / 41s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: A- Clarity: B+ Strength: B	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B

9	"The Arctic states are completed and integrated by customary international sea law and several treaties. The Arctic includes areas of national sovereignty. Cuba fully respects these sovereign rights and will be ready to play their role to confront global challenges with its scientific and technological expertise and leading companies to contribute to a sustainable Arctic development while respecting the ecosystem and indigenous people. In this frame, Cuba expects the extending bilateral and multilateral cooperation in the Arctic, in the framework of international law to create a committee to cooperate in the following subjects: wider access of geo-strategic interest, scientific cooperation on climate change and environmental protection, economic expertise in the use of mineral resources, cooperation on human lives in extreme climate conditions." (129 words / 38s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B+ Clarity: B- Strength: B-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
10	"The Dominican Republic says we have to take care of the arctic. It is important because there are many resources there which the world will be sad to lose. It also has endangered species. We have not been careful with the arctic. Countries could also go to war over it. We need to be responsible. This means setting up mechanisms to protect the environment. We should also create treaties to protect nature and the natural resources there. We owe it to our children and their children to take care of the arctic. We will have failed them if they ask us "why were you not responsible". We need to be able to have an answer. As ECOFIN, we need to make sure the United Nations ensures that its member states use the arctic responsibly or do not use it at all. If we want an arctic when we are older we need to protect it today!" (156 words / 46s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B+ Clarity: A- Strength: B	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
11	"Honorable chair, distinguished delegates: 800 million people across the globe are living without access to clean water. Half of them live in Sub-Saharan Africa. The Kingdom of Norway believes that our efforts should focus there, and these efforts should involve the entire international community. Norway has donated over $975 million in aid to sub Saharan countries and invites countries who have not donated to join the effort. Norway proposes we focus on community-oriented aid in the form of water harvesting devices. A good option is the water harvester developed by the University of Akron in Ohio which produce up to 10 gallons of drinking water per hour from thin air. To ensure long term success, the United Nations should transfer the funding after there is sufficient training to build and operate the water-gathering device without continued external support. Community-oriented aid is a big step towards universal access to clean water. Let's say H2"NO" to the water crisis!!!" (159 words / 46s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B- Clarity: A- Strength: A	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
12	"Honorable Chair, Distinguished Delegates, Comoros agrees with Canada, Ecuador, and Tanzania that cleaning water is the top priority and disagrees with Norway and Chad that water creation is an affordable or viable solution. In small countries like Comoros the largest problem is water so polluted it is fatal to drink. Karthala, our most active volcano, erupted in April 2005 and May 2006. Each time, it destroyed multiple freshwater sources. It also took years to clean the volcanic ash that comes back as acid rain. Our salvation was UNICEF and the European Commission's Humanitarian Aid Office which give us $1.3 million to clean more than 1,500 reservoirs and bring fresh water to more than 150,000 people. Comoros is aware that most countries do not have active volcanoes. However, water pollution is a global phenomenon, whether from other natural disasters or corporations who pollute without accountability. Comoros suggest the creation of UN-funded permanent water filtration programs. These can also be funded with proceeds from Canada's policy to make polluting companies pay for their own cleanup." (176 words / 51s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B+ Clarity: A- Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
13	"Most esteemed chair, fellow delegates. The post-colonial era has seen the independence of many of the poorest countries in the world, Chad among them. However, Chad strongly believes that this independence should not have to result in a new dependence upon foreign aid for survival. Because of this, Chad believes that the future for Africa rests on multilateral efforts within the UN to assist in the global effort to solve the pressing environmental, social and political crises that face both Chad and the rest of Africa. This process started a while back and it is beginning to bear fruit with many of the African countries emerging from the problems of recent years. The result of this is that now Chad, as an African nation, can work together with the rest of the continent to help solve the major issues of today such as the problems of the rights of the child, combating diseases such as Ebola, which kill thousands, and helping African nations achieve the permanent seats which they deserve in the Security Council. It is with these principles and these objectives held firmly in mind that the Chadian delegation enters this conference. Honorable President, fellow delegates, thank you very much." (201 words / 59s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: A- Clarity: B+ Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
14	"Fellow delegates, respected Chairs and esteemed President, As members of the same world, it is only fair that each one of us work hard and together to improve the peace, growth and prosperity of the entire human race. In Colombia alone, almost 5,000 children under the age of 5 are set to die this year due to malnourishment, illness, abuse and a general lack of both national and international assistance and aid. But this is not a problem only affecting Colombia and other developing nations. Take the country we are in right now, for example. In the Netherlands, an estimated 1,000 children are set to die for the exact same reason. But why, why, delegates, why is it that we are letting children around the world die? The answer lies in the insufficiency of international cooperation to help humanity. This is something that is not only needed to help protect child rights, but also to help take nations from a state of relief to a state of development after natural disasters and to help with a true establishment of a culture of peace. Thank you." (187 words / 54s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: B+ Clarity: A- Strength: A	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
15	"Honorable Chair, fellow delegates, A very special thing has happened today. Observe how many ladies sit among the men as equals. They're not in the fields, they are not instruments of pleasure available for purchase. They're here with a voice that will be respected. The United Nations Development fund for Women could not be more proud of the progress we have made. Over the years, UNIFEM has worked to give a voice to those who are not allowed to speak. Women everywhere were once denied the right to life, to identity, to self-worth. But now, as the equal presence of women in this forum clearly screams out, women are getting status on society's ladder. UNIFEM has given women a chance to voice their opinions. UNIFEM has fought to ensure women can enjoy rights in issues such as medication and health care, along with maternal care and freedom and protection from violence. This week, it'll be the responsibility of this forum to continue with this revolution. This week, don't just be a voice for a delegation, be a voice for all the women who have still not been allowed to find their own. Thank you." (194 words / 56s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: A- Clarity: B+ Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A Brevity: A Clarity: A Strength: A
16	"Honorable chair, fellow delegates and distinguished guests: The delegation of Sudan realizes the importance of a world without a digital divide between MDCs and LDCs, genetic privacy and nondiscrimination, as well as resolution on the issues between Israel and Palestine. However the delegate of Sudan wishes to	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback

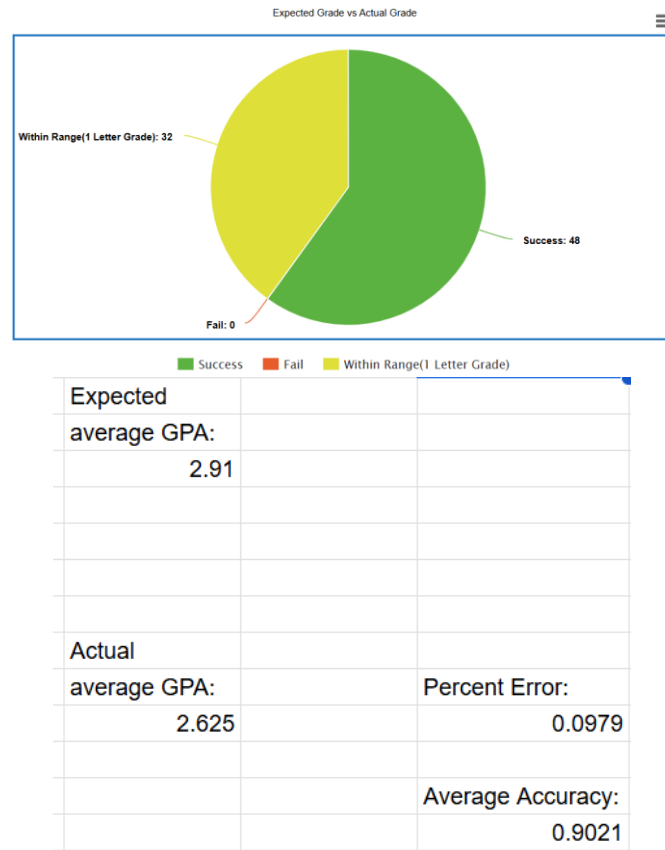|			
	especially focus on digital access among all countries, for all people. The digital divide between MDCs and LDCs is something that needs to be bridged in order to help lower the rates of poverty worldwide and give many LDCs a better economy. As an LDC, Sudan is affected by this issue because of its lack of up to date digital access. The delegate of Sudan believes that a possible solution to the bridging of the digital divide is assistance from MDCs to LDCs in the form of programs, donations, and training. The delegate of Sudan looks forward to hearing and debating other delegations' opinions on this and other issues." (156 words / 45s)	on how the speech can be improved. Overall: B+ Brevity: B+ Clarity: A- Strength: B+	on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
17	"Honorable chair and fellow delegates; the United States is extremely privileged to attend this conference. In the past few years, the United States and several allies have been fighting a campaign in Iraq as part of its war on terror. Many lives were lost, but after long, hard combat the US finally defeated dictator Saddam Hussein. After his removal, the United States has been working to rebuild Iraq into a democracy and soldiers have stayed behind to eliminate insurgents. However, the delegation from the United States feels that its progress in Iraq has come to the point where the Iraqi government can handle itself independently, as is needed for any strong democracy. Beginning in 2011, the United States will begin withdrawing its soldiers from its Iraq. It has faith in the Iraqi people and know that they will be able to become a healthy nation. Thank you." (147 words / 43s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: B- Clarity: A- Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
18	"*Honorable chair, fellow delegates and distinguished guests,* The United States of America is grateful for the opportunity to be a part of this conference and the chance to find suitable solutions for environmental issues, such as nuclear waste disposal. Nuclear waste disposal is an important issue to the modern world because nuclear power greatly benefits human society, while at the same time potentially resulting in serious consequences. Exposure of radioactive waste in the environment must be eliminated if nuclear power is to be used, and better disposing methods of nuclear waste must be achieved. The United States highly values the environment and its well being. Therefore, it looks forward to discussing this and other issues in the following days. Thank you." (129 words / 38s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: A- Clarity: A- Strength: A-	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: A Strength: B
19	"Honorable Chair and fellow delegates and distinguished guests, thank you for according us your time to deliberate this speech which Morocco believes is capital. Despite many efforts made by Morocco, gender equality in access to primary and secondary education is essential for a developing country such as Morocco. Women in Morocco usually receive less education than men. The data shows that 40% of women and 60% of men in Morocco receive a basic education which has a huge impact on the society. Morocco has already made considerable effort to enforce the laws but still has some progress to make. Morocco is looking forward for the support of UN countries beyond the borders for solutions to this International issue." (119 words / 35s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: A- Brevity: A- Clarity: A- Strength: B+	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: B Strength: B
20	"Your Excellency the Secretary General, Mr President, honorable delegates, Italy has been in the UN since 1955 and is the 5th top contributor to the UN budget, as well as being a founding member of the E.U. It administers justice, and is firmly committed to fighting crime, racial and religious discrimination child slavery and the effects of natural disasters. Italy believes in a pragmatic approach to the world's problems such as conflicts in the Balkan States, The question of Kashmir,, hazardous waste, sexually transmitted diseases anti-personnel mines and others that will be discusse d during our meeting in this wonderful college. Italy wishes you all the most interesting debates during these two days! Thank you. I yield the floor to the President." (122 words / 36s)	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B+ Brevity: A- Clarity: B+ Strength: B+	ChatGPT provides overall grade and a grade for each criteria it was asked. Feedback for each criteria is also provided. ChatGPT also appends feedback on how the speech can be improved. Overall: B Brevity: B Clarity: BStrength: B

Figure 7.Table of experiment

Figure 8. Expected grade vs actual grade

Out of 80 outputs, 48 were successful while 32 managed to stay within range. With our schema for accuracy, we got 0 test cases that had failures. The overall accuracy of our system is roughly 90.21% and a percent error of 9.79%. ChatGPT managed to always provide proper formatting for the output in layout that was expected. However a consistent grade difference of one letter was observed throughout multiple times. This inconsistency in grading contributed to the amount of within range outputs. The lack of letter grade signs in the ChatGPT grade outputs compared to the expected outputs resulted in the letter grade difference. In the future, I will have to modify the prompt given to ChatGPT to encourage it to include letter grade signs.

5. RELATED WORK

The paper Enhancing Public Speaking Skills Through AI-Powered Analysis and Feedback by Soham Padia et al. focused on utilizing AI in order to enhance public speaking skills similar to our project [11]. The difference is that additional technologies such as facial expression analysis using OpenCV and YOLO are incorporated in order to better analyze a user's speech. We believe that while this approach is more accurate, our approach is better suited to general purpose use and is more accessible to the general user. Combining AI analysis alongside a VR environment allows the user to better simulate a real life speaking scenario.

The paper Virtual Reality Public Speaking Training: Effectiveness and User Technology Acceptance by Manual Bachmann et al. attempts to gauge the effectiveness of what they refer to as Virtual Reality Public Speaking Training (also known as VRPST) [12]. Their findings demonstrated that users using VR public speaking training methodologies had better skills at the

end of their sessions than those who prepared individually. However, their approach does not incorporate AI systems, and only attempts to simulate environments. We believe that our incorporation of a large language model into our testing allows for the VRPST that the researchers mention to be condensed and accessible to just singular users.

The paper SpeechMirror: A Multimodal Visual Analytics System for Personalized Reflection of Online Public Speaking Effectiveness by Zeyuan Huang et al. focuses on utilizing computer vision in order to evaluate public speaking skills [13]. It aims to use visual techniques to provide users insights on their public speaking proficiency. We believe that our approach of using virtual reality alongside immediate artificial intelligence based analysis instead of only visual analysis is a better methodology that individuals can use to grade their own performance. Simulating public speaking environments using virtual reality allows users to simulate a more immersive experience and better practice their skills.

6. CONCLUSIONS

While we believe our project consistently reaches our initial goals, there are also limitations that would be pertinent to address. For instance, the only data we provide to the LLM model for analysis would be a transcript of the user's speech, which is done via OpenAI's Whisper library [14]. While this is consistent in its accuracy, it is not completely error prone. We believe that in the future we will need to add other data to introduce redundancy, such as the pure audio file and find some way for the LLM to process it, perhaps by switching to one that supports multimodal support. Additionally, the environment set up for the application is suitable enough for someone to give a presentation in, however more immersive graphical fidelity for the environment would make the scene more believable and allow users to be more comfortable using the application as a means to practice. Lastly, our application would also benefit from having users be able to somehow import their presentation materials such as a slideshow or speaking notes, which would again make the experience more true to life and make users more comfortable giving presentations to a virtual environment[15].

REFERENCES

[1]   Parvis, Leo F. "The importance of communication and public-speaking skills." Journal of Environmental Health 63.9 (2001): 44-44.

[2]   Anthes, Christoph, et al. "State of the art of virtual reality technology." 2016 IEEE aerospace conference. IEEE, 2016.

[3]   Hadi, Muhammad Usman, et al. "A survey on large language models: Applications, challenges, limitations, and practical usage." Authorea Preprints (2023).

[4]   Chaturvedi, Anurag, Sofia Singh, and Jitendra Singh Jadon. "AI-Powered Emotion Recognition: Exhibiting Human Emotions Through Text, Speech, and Facial Expressions." 2025 International Conference on Engineering, Technology & Management (ICETM). IEEE, 2025.

[5]   Chen, Juntong, et al. "InterChat: Enhancing Generative Visual Analytics using Multimodal Interactions." Computer Graphics Forum. 2025.

[6]   Kasneci, Enkelejda, et al. "ChatGPT for good? On opportunities and challenges of large language models for education." Learning and individual differences 103 (2023): 102274.

[7]   Welsby, Philip, and Bernard MY Cheung. "ChatGPT." Postgraduate Medical Journal 99.1176 (2023): 1047-1048.

[8]   Chowdhary, KR1442. "Natural language processing." Fundamentals of artificial intelligence (2020): 603-649.

[9]   Batista, Josué R. Learn OpenAI Whisper: Transform your understanding of GenAI through robust and accurate speech processing solutions. Packt Publishing Ltd, 2024.

[10]   Xu, Canwen, et al. "Baize: An open-source chat model with parameter-efficient tuning on self-chat data." arXiv preprint arXiv:2304.01196 (2023).

[11] Cherner, Todd, et al. "AI-powered presentation platforms for improving public speaking skills: Takeaways and suggestions for improvement." Journal of Interactive Learning Research 34.2 (2023): 339-367.

[12] Bachmann, Manuel, et al. "Virtual reality public speaking training: effectiveness and user technology acceptance." Frontiers in virtual reality 4 (2023): 1242544.

[13] Huang, Zeyuan, et al. "SpeechMirror: A multimodal visual analytics system for personalized reflection of online public speaking effectiveness." IEEE Transactions on Visualization and Computer Graphics 30.1 (2023): 606-616.

[14] Alberts, Ian L., et al. "Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be?." European journal of nuclear medicine and molecular imaging 50.6 (2023): 1549-1552.

[15] Van Ginkel, Stan, et al. "Fostering oral presentation competence through a virtual reality-based task for delivering feedback." Computers & Education 134 (2019): 78-97.