iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XI **Month of publication:** November 2021

DOI: https://doi.org/10.22214/ijraset.2021.38917

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XI Nov 2021- Available at www.ijraset.com

Virtual Voice Assistant Based on Voice Flow

Dhruv Piyush Parik[1], Bhavuk Mittal[2], Amogh Pal[3]
[2, 3]*School of Computer Science Engineering, Vellore Institute of Technology, Chennai, India*
[1]*School of Electronics and Communication Engineering, Vellore Institute of Technology, Chennai, India*

Abstract: ***We face a perennial pandemic that forces everyone to stay within their premises, which engenders a decline in social interaction between individuals. Moreover, some people fear missing out, which correlates to the fact that individual proclivity towards human interaction is drained, which increases symptoms of depression and raises the bar for anxiety—inspiring from such social circumstances, we want to develop a system that aids in mitigating these social problems. Our system has an artificial neural network layout that enhances personalisation through voice application. An Autonomous Virtual Assistant System is an effective method that shall help a person deprived of social interaction get engaged in a gregarious task, deal with various problems faced by introverts by reducing the psychological impact of COVID 19.***
Keywords: Google Assistant, Intergromat, APIs, Voiceflow, Speech to text

I. INTRODUCTION

Human psychology is shaped so that when we get rattled, we are steered to share it with a person. According to a study by Jonah Berger, our autonomic nerve is activated. We are carried away to convey the message to anyone around us. However, during the global pandemic, humans got isolated, and there was no one left to talk to, which led to a rise in depression and anxiety.

In recent years, the capability of machine learning has been added to voice recognition software, resulting in a level of 'intelligence' that was previously not prominent to the masses. Google released the Voice Search App for the iPhone in 2008, and Siri was released in 2011, providing consumers with their own digital personal assistant.

Voice recognition enabled consumers to handle their gadgets more efficiently than ever before, and we are going to utilise this in our research to curb the effects of the global pandemic, which is disturbing mental health in most countries; youths have seen a rise in suicidal intentions, people have economic crises that lead to depression and anxiety. So, we are concentrating on a system that will help people share anything they want to share with our voice application to continue the conversation with them, which intrigues their interest.

An Autonomous Virtual Assistant System is a powerful system, which shall help to carry the conversation by trying to understand their problems and help them in day-to-day activities.

II. METHODOLOGY

People are afraid to talk to or share their feelings with others. Individuals frequently fail to locate the perfect person with whom to share their experiences. People will not hesitate to communicate their sentiments with the aid of our basic application, and they will be able to do it at their convenience and at their own pace.

A. Research Approach

Our research paper revolves around human emotions and sentiments by lifting the burden of therapy from afflicted people. This voice application is just like an all-time social companion ready to talk according to your convenience. The study also looks at the user experience of using the application, finding that most users are satisfied with the User Interface of Voice Application.

B. Method Approach

Our voice application system processes the user's voice, and then our model interlinks to Live APIs to provide the user desired results by utilising voice flow models.

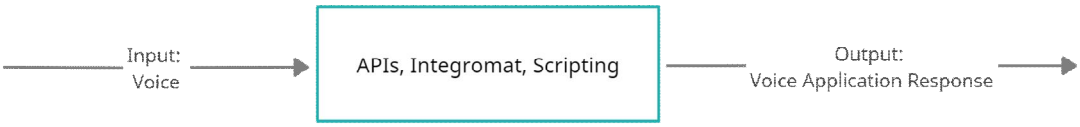


Fig. 1 Block Diagram

©IJRASET: All Rights are Reserved

C. Voice Flow

Voiceflow is a cutting-edge creative tool for teams focused on conversation design and product development. Voiceflow streamlines the design and prototype process while facilitating design-dev handoff and live experience deployment. The helper speaks as a result of the Speak step. We can add audio and SSML effects and enter what we want the assistant to say in numerous voices. We can even modify the assistant's voice and make the prompts random. We can use the Audio Step to play and stream audio files less than 240 seconds long. Audio can be supplied by uploading it directly to Voiceflow or providing the audio file's stream URL. To make our audio streams more dynamic, we can utilise a variable to host the stream connection.

When creating for Amazon Alexa and Google Assistant channels, the Visuals Step displays visuals both in the prototyping view on the Custom Assistant channel and actual devices.

D. API

API is for Application Programming Interface, and it is a software mediator that allows two apps to communicate with one another. We utilise an API every time we use an app like Facebook, send an instant message, or check the weather on our phone. A REST API is an API that follows the REST, or representational state transfer, architectural style's design principles. REST APIs are sometimes referred to as RESTful APIs because of this. We used REST API to get and push responses in several modules for our investigation.

E. Integromat

Intergromat is a tool that aids in the automation of manual activities such as handling various APIs. Integromat is a programme that allows you to automate manual operations without the need for coding. They claim to themselves as "the internet's glue," assisting their clients in connecting apps and services. Instead of requiring coding, Integromat has simplified complex APIs into simple pieces that can be linked with a simple drag-and-drop. Integromat refers to these blocks as 'Modules,' and the visual links between them as 'Scenarios.' We can quickly redeem information from the APIs we have utilised with the aid of Integromat. This makes our voice application easy to create and use.

Fig. 2 Workflow of Integromat via webhooks

©IJRASET: All Rights are Reserved

III. ALGORITHM

This section discusses the analogies we utilised in our voice application. Here, we've gone through in detail how our application works and how we respond to it.

A. Voice Flow

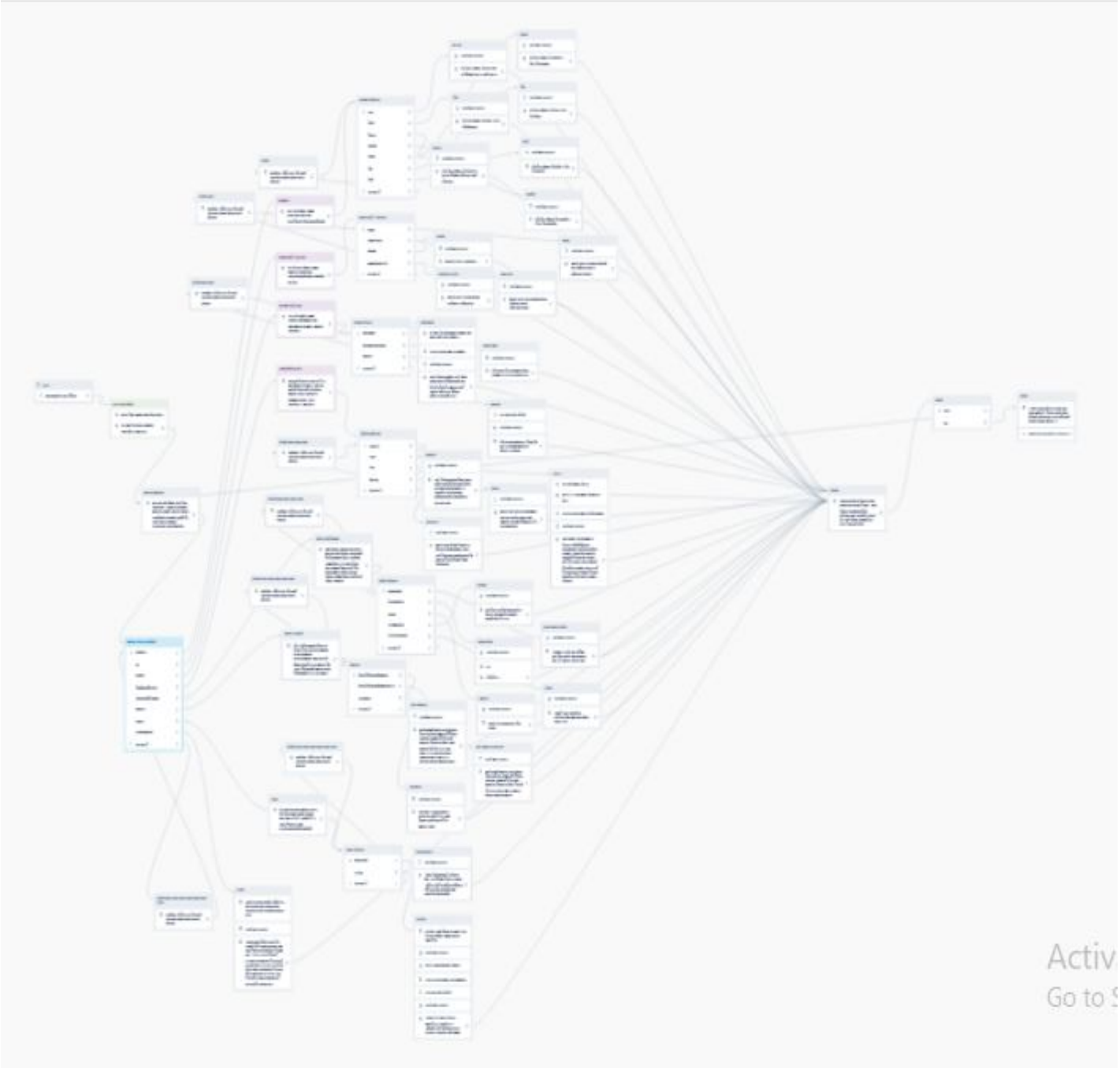


Fig. 3 Virtual Voice Assistant based on Voiceflow

Voice flow is the web application to build powerful voice applications. We have used voice flow to make virtual assistant. It is based on a simple interaction between the speech system and the user. We created a prototype that combines many sample conversations as shown in fig. 3, we have used responses such as speak and dialog box, user inputs by choice and prompt, several APIs, and Intergromat.

©IJRASET: All Rights are Reserved

B. Response

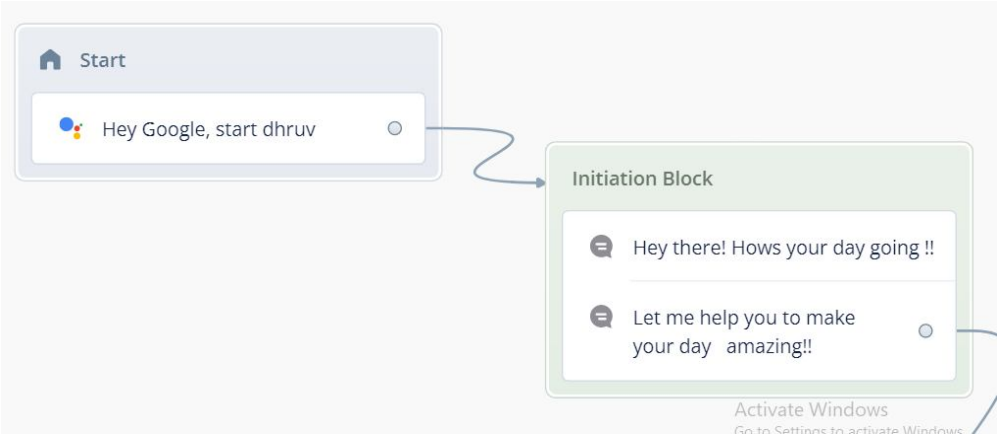


Fig. 4 Speak Module Representation

At various phases of our programme, we employed numerous responses to make our voice application interactive: Speak, Catch Audio, and Display Cards. At first, the user is pleasantly greeted as "Hey there! How is your day going? Let me help you to make it amazing?" Messages like this are pre-defined in voice flow's speaking module. The message is shown automatically with the aid of cards.

For example, if you try to converse with the app, your audio is instantly collected by the catch block, and the cards respond meaningfully.

C. User Input

We have implemented choice and prompt blocks to take input from users. We have different choices for a particular block like if you want to talk about animals, our voice application will ask if you are about which animal you want to about, it will list several names of animals with helpful prompts, and the user will be able to make a choice.

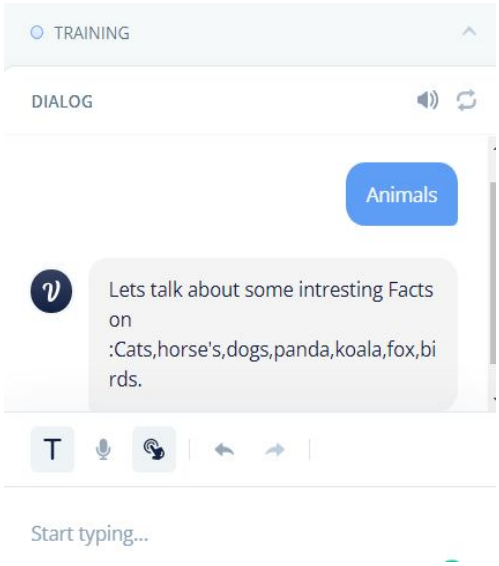


Fig. 5 User Input in Voice flow

D. Logic

Capture, condition, and leave events govern the logic portion. The Condition Step, often known as the IF Step, enables you to direct the user down different conversation paths based on the value of variable conditions. For example, if the user wants to learn more about animals, he or she will be sent to a page that has information about various species. If the user wants to learn about movies, the system takes them along a different path. The Capture Step records and saves the user's comment in a variable. The Exit Step will bring the discussion and app to a complete halt, with no more messages being sent to the user. It will be considered as the end of conversation.

©IJRASET: All Rights are Reserved

E. Intergromate

To retrieve data from the API, we utilised Intergromate and some custom code.

There are two types of API in our Voice Application:

1) Live API

2) Rest API

We have used Live API at blocks where we need live information like Latest News.

For example, In the news section, we have NY Times Live API to fetch current news that keeps getting updated in real-time. We utilised Rest API in areas where we needed information or facts because there is a pre-set database for facts and information, such as when we needed to share a fact about the dog or how to acquire a poem for our user.

IV. RESULT

People experienced a great deal of mental stress and anxiety during this period of the global pandemic. When they were restricted in their houses, it was frequently difficult for them to locate someone with whom to converse. In many situations, it led to social anxiety disorder.

We considered developing a voice application to remove the barrier in people's thinking about who to speak with. They may just engage in a one-on-one chat about nearly anything, from the weather to sports, jokes to movies and whatnot. We made an effort to include people of all ages and to assist them in becoming more participatory and livelier.

We have made a voice application to make communication easier for people during this global pandemic. Voice flow has been used to make our voice application efficient which is capable to talk to people especially when there is no one to talk around. We have used by voice recognition, text to speech, and APIs in our voice application to reach some conclusions for our research paper.

We used multiple APIs to ask a query and then respond based on the response.

If our user is teenager, he will be curious about what is going on around him. For instance, suppose he asks the voice app a question concerning animals. The API responds by informing him/her of some fascinating information about certain animals. According to the user's response, the data about the specific animal are automatically retrieved from the REST API.

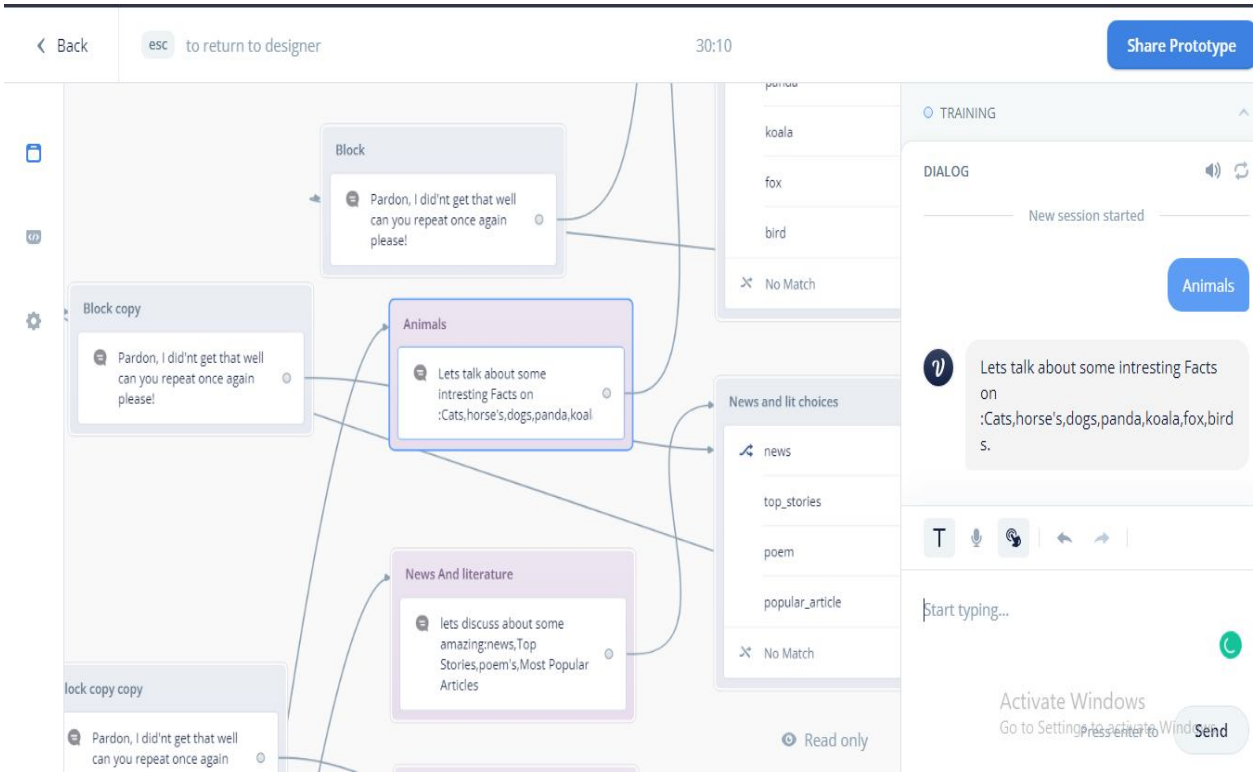


Fig. 6 Output when user want to talk about animals

If the user is a young adult who wants to discuss drinks. The food and drinks API will respond by telling you what his favorite drink is. It will try to carry on the discussion like a regular person and will make you feel as at ease as possible.

©IJRASET: All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XI Nov 2021- Available at www.ijraset.com

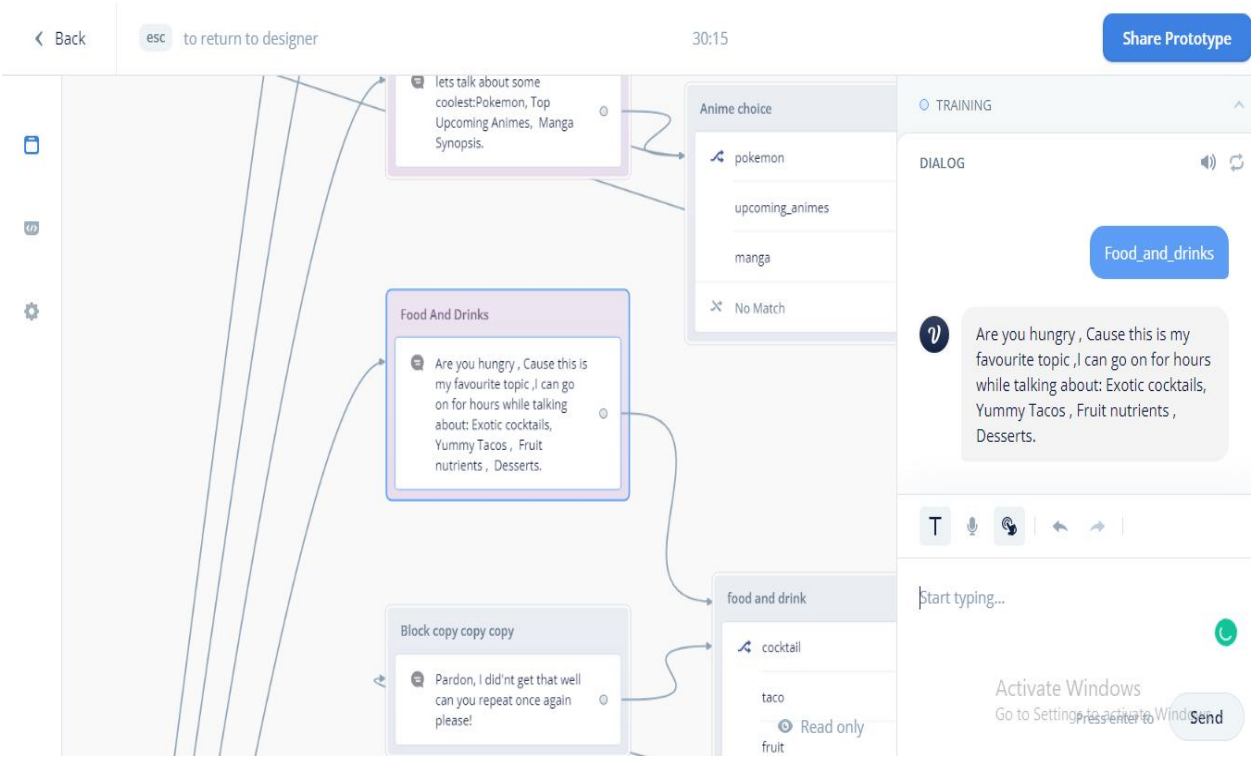


Fig. 7 Output when user want to talk about foods and drinks

If he is a middle-aged man who likes to hear the latest news. The News API will retrieve the news from the NY Times Live API and keep you up to date on the top stories that are trending.

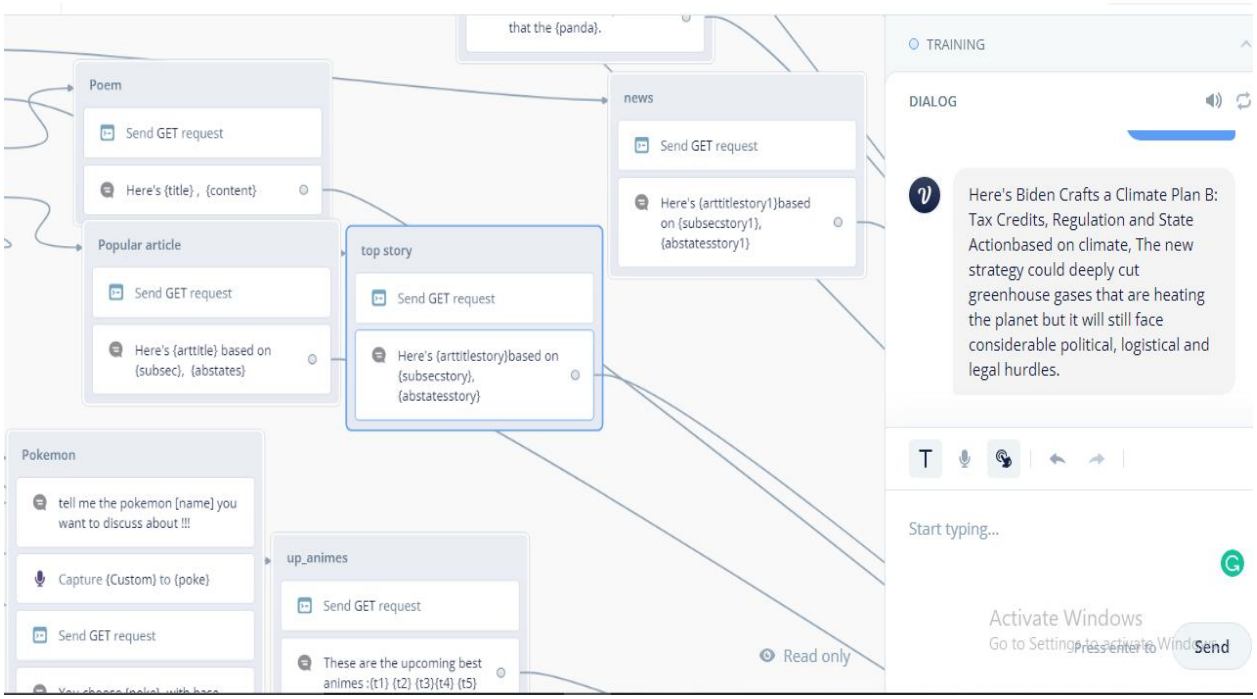


Fig. 8 Output when user want to talk about News

If an elderly guy has nothing else to do with his life, he might just request to listen to poetry from the voiceflow. The News API block will read a poem of his choosing at random.

©IJRASET: All Rights are Reserved

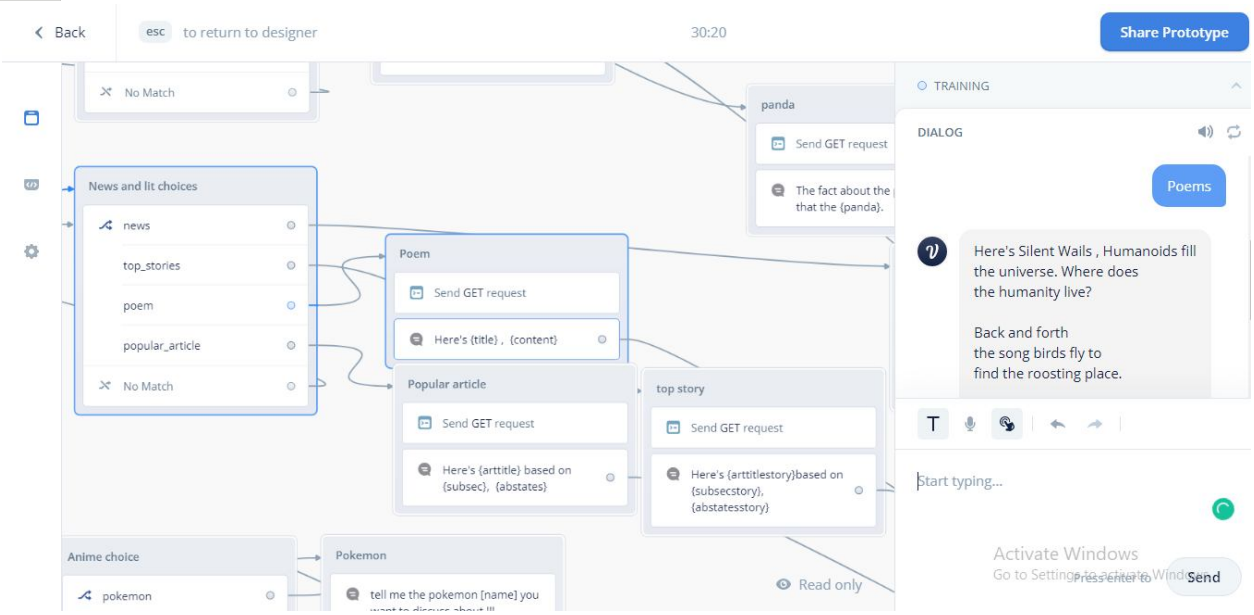


Fig. 9 Output when user want to talk about Poems

Voice recognition is an in-built feature of voiceflow that makes it so simple to use. We recognize what users say and give responses according to the situation to carry out an interactive conversation. By understanding the sentiments of users and presenting predefined responses to them, they tend to feel conversing with a human rather than a voice application.

V. CONCLUSION AND FUTURE SCOPE

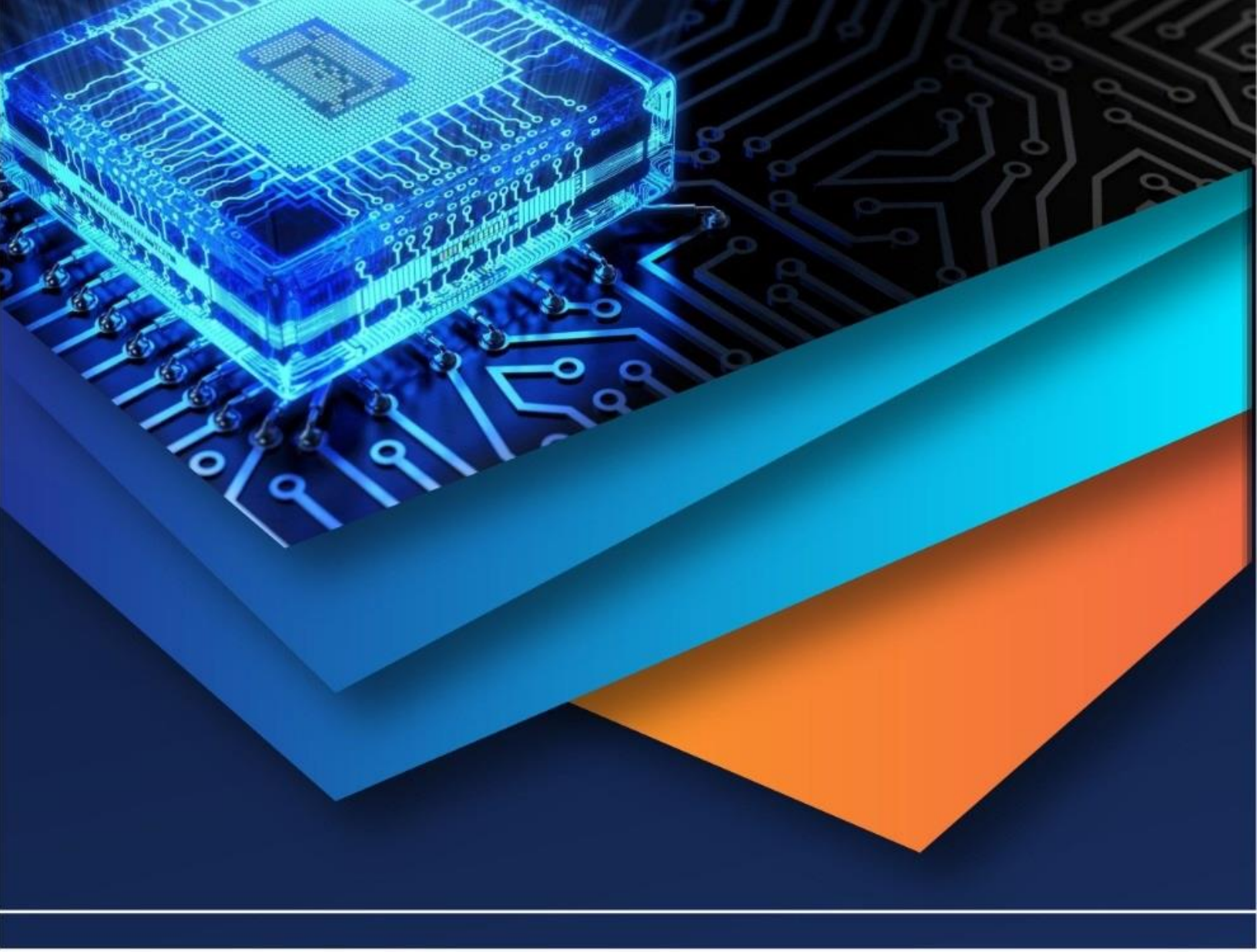
This project's work scope includes interacting with a wide range of persons who suffer from social anxiety disorder. People frequently do not perceive this to be an illness. As a result, they frequently fail to come up with a solution. Humans are social beings. They can't exist without engaging with their environment. At the same time, it is not their responsibility if they lack the ability to initiate a discussion. In the shape of our voice application, we have created a wonderful answer to such an unsolvable problem. It crosses the borders of people's minds about who and how to talk to. We may just speak to our voice app about anything at any moment. According to studies, certain disorders do not have medications. They can be treated if the proper steps are taken. This is a slogan that even psychologists adhere to. They advise their patients to express themselves and talk about their feelings rather than burying them and suffocating them.

Taking this a step further, our voice program may be installed in clinics and utilized for therapy. It may also be put up at old age homes and orphanages, where individuals typically have no one to talk to. Thus, our application will aid in reaching out to the underserved segments of society or those with whom we frequently avoid conversing.

REFERENCES

[1] Sarah K. Schäfer, M. Roxanne Sopp, Christian G. Schanz, Marlene Staginnus, Anja S. Göritz, and Tanja Michael, Impact of COVID-19 on Public Mental Health and the Buffering Effect of a Sense of Coherence, Psychother Psychosom 2020; 89:386–392.
[2] Gisli H. Gudjonsson, Jon Fridrik Sigurdsson, Olafur O. Bragason, Emil Einarsson, and Eva B. Valdimarsdottir, Compliance and personality: the vulnerability of the unstable introvert, European Journal of Personality Volume: 18 issue: 5, page(s): 435-443.
[3] Tiia Tulviste, Luule Mizeraa and Boel De Geerb, "There is nothing bad in being talkative": Meanings of talkativeness in Estonian and Swedish adolescents, Journal of Pragmatics Volume 43, Issue 6, May 2011, Pages 1603-1609.
[4] Candace Kamm, User interfaces for voice applications, Colloquium Paper Vol. 92, pp. 10031-10037, October 1995.
[5] Jay G. Wilpon, Voice-processing technologies Their application in telecommunications, Colloquium Paper Vol. 92, pp. 9991-9998, October 1995.
[6] Bishnu S. Atal, Speech technology in 2001: New research directions, Colloquium Paper Vol. 92, pp. 10046-10051, October 1995.
[7] Nirmita Panchal , Rabah Kamal , Cynthia Cox, and Rachel Garfield, The Implications of COVID-19 for Mental Health and Substance Use. [Online]. Available: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)