iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42193

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: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

AI System With Voice Modulation Using Neural Network

M. Roshini[1], S. Nandhika[2], S. Prathipa[3], Dr. Rajasekar Velswamy[4]
[1, 2, 3, 4]*Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*

Abstract: ***A virtual assistant, also called AI assistant or digital assistant, is an application that understands voice commands and completes tasks for the user. Popular virtual assistants currently include Amazon Alexa, Apple, Siri, Google Assistant. In our project we'll create an AI system using Neural network algorithm and NLP Techniques. AI System which is able to perform all the actions-speaking, listening task, google search engine tasks, playing YouTube videos, tracking current location, date, time, day, year etc. Additionally, to these features AI Assistant will recognize emotions and communicate back to the speaker ,changing the voice of the assistant and recognize other language.***

I. INTRODUCTION

Artificial intelligence (AI) the flexibility of a information processing system or computer-controlled robot to perform tasks commonly related to intelligent beings. Such tasks, historically performed by a private assistant or secretary, include taking dictation, reading text or email messages aloud, looking up phone numbers, scheduling, placing phone calls and reminding the end user about appointments. An AI system can perform tasks or services for private supported command,questions. Popular virtual assistants currently include Amazon Alexa, Apple's Siri, Google Assistant. The existing popular AI Assistant is Amazon Alexa. It starts with signal processing, which provides Alexa as many chances as possible to create sense of the audio by cleaning the signal. the thought is to enhance the target signal, which implies having the ability to spot ambient noise just like the TV and minimize them. To resolve these issues, seven microphones are wont to identify where the signal is coming from therefore the device can specialize in it. the following task is “Wake Word Detection”. It determines whether the user says one among the words the device is programmed to wish to show on, like “Alexa”. this can be needed to reduce false positives and false negatives. If the wake word is detected, the signal is then sent to the speech recognition software within the cloud, which takes the audio and converts it to text format.

When users say ‘Alexa’ which wakes up the device. The wake word put the Alexa into the listening mode and prepared to require instructions Invocation name is that the keyword wont to trigger a selected “skill”. Users can combine the invocation name with an action, command or question. All the custom skills must have an invocation name to begin it.

‘Taurus’ is an utterance. Utterances are phrases the users will use when making a call for participation to Alexa. Alexa identifies the user’s intent from the given utterance and responds accordingly. So basically, the utterance decides what user want Alexa to perform.

After that Alexa enabled devices sends the user’s instruction to a cloud-based service called Alexa Voice Service (AVS). Alexa Voice Service is that the brain of Alexa enabled devices and perform all the complex operations like Automatic Speech Recognition (ASR) and tongue Understanding (NLU). Alexa Voice Service process the response and identify the user’s intent, then it makes the online service request to 3rd party server if needed.

Our AI system also can perform all the tasks like Google program, search YouTube video , tracking locations ,playing music, answering to the questions and together with these features our Assistant can even recognize the emotion of an individual by analyzing the voice of an individual and may answer accordingly.

In our project we have used four modules

1) Base
2) Emotions
3) Features
4) Advanced features

In today’s world most of the people are much involved in their day-to-day work due to this reason people cannot find time to relax themselves. Even some people work fraught they don’t have time to speak to others and to release their stress. So, our AI System are more helpful for people to speak with this AI whenever they're under stress or pressure because the AI can recognize their emotions and respond back during a friendly way.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

II. RELATED WORK

In our project we have 4 modules:

A. Creating Base

1) Speak

AI Assistant is trained in such a way that it reads the given text within the variety of speech.
Packages used: pyttsx3

2) Listen

AI Assistant is trained in such a way that it takes input as speech of a user using microphone and converts the speech into text.
Packages used: Speak Recognition

3) Brain

In this brain of AI Asssistant is developed using neural network where user input is taken within the type of input layer and processed together with hidden layer and after performing the operations inside the neural network the required result's given using output layer.
Packages used: Torch

4) Neural Network

In this we perform NLP techniques the primary one is tokenization where a sentence is abate into set of words called token for better understanding of every word then it follows stem technique where the assistant takes only the basis word and finds the precise meaning of it by comparing with the intents then finally it follows bag of words where we'll be splitting each word within the sentences and adding it to an array. which is able to initially be a listing of zeros with the dimensions adequate the length of the all words array. If we've a array of sentences = ["hello", "how", "are", "you"] and an array of total words = ["hi", "hello", "I", "you", "bye", "thank", "cool"] then its bag of words array are going to be bog = [0 , 1 , 0 , 1 , 0 , 0 , 0]. we'll loop over the each word within the all words array and also the array resembling each word. If a word from the sentence is found within the all words array, 1 are replaced at that index/position within the array.Neural network is performs these tasks inside and provides the output.
Packages used: NumPy, NLTK

5) Intents

Created intents.json file inside this we've tags, patterns and responses. If a user says hello it goes inside greetings tag and checks for patterns if it matches then the response from AI is given as hello. Like wise we've created multiple intents like bye, health, date, time, day, Wikipedia, google, location, temperature , current status.

6) Train

Firstly, we'll be loading the intents.json file then will will separate all the tags & words into their separate lists then we'll be cleaning the info by implementing the functions that we created earlier . subsequently we transform the information into a format that our PyTorch Model . Every Neural network features a set of hyper parameters that require to be set before use.Before Instantiating our Neural Net Class or Model, we are going to first defined some hyper parameters which might be changed accordingly. We used __getitem__ and __getitem__ magic funtions. Then we Instantiated the model, loss and optimizer functions. subsequently we trained the model then the training was completed and file saved to data.pth. So during this module we trained the AI using intents.
Modules used : json ,numpy, torch

7) Creating the AI

In this we Loaded our Saved Model. Finally Our AI Assistant training was complete. Then we Trained the AI to perform all the functions : Listening, speaking and performing all the actions as per the request of the user.
Modules used: random

B. Emotions

Created more intents to acknowledge the sentences of the user.
Here additional tags are created like identity, distinguishing the emotions by creating new tags for various patterns and created responses.
So that the AI can understand exactly what the speaker is asking or talking about.

©IJRASET: All Rights are Reserved

C. Features

AI performs tasks requested by the speaker:

(i)Non –Input:

In this the AI assistant will perform time , day, date task.

For example if a user asks for date then the AI will respond and display the date of the present day.

(ii) Input:

In this the AI assistant will perform tasks based on particular question of the user. For example if a person asks a question what is environmental effects then the AI will answer to the question by saying the definition, types etc.Various tasks performed are: Answering to the questions ,Wikipedia search, Google search engine tasks ,YouTube search ,Pointing out the location , Performing mathematical operations.

D. Advanced Features

We trained our AI to Recognizing other languages.

Also Changing the voice of the AI Assistant.

In this paper we've used NLP Algorithm .It handles interactions between machines and natural languages of humans within which computers are entailed to analyse, understand, alter, or generate natural language. NLP scans a sentence from left to right to analyse the meaning of a sentence.

In this paper we Perform a number of the methods of Natural Language processing tasks like

Tokenization:

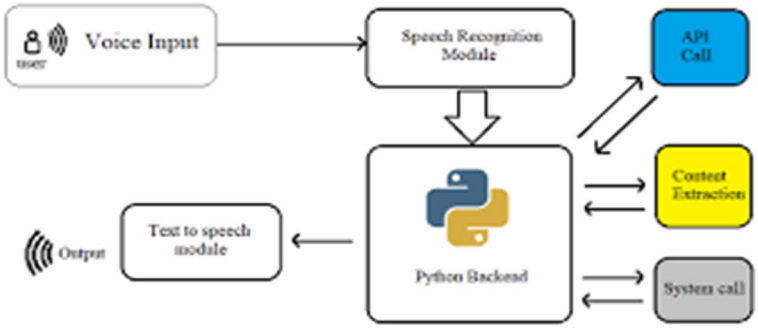
Tokenization is breaking the raw text into small chunks. Tokenization breaks the raw text into words, sentences called tokens. These tokens help in understanding the context or developing the model for the NLP. The tokenization helps in interpreting the meaning of the text by analysing the sequence of the words.

Stem: Stemming is basically removing the suffix from a word and reduce it to its stem. the most aim is to scale back the inflectional sorts of each word into a standard base word or descriptor or stem word.

bag_of_words

After processing through tokenization and stem bag_of_Words is being processed where we get the desired results and also the assistant performs the desired result.

III. RESULTS AND DISCUSSION:



In this paper we have first taken the input as speech of a user for this we have used microscope so that only the voice of a person can be taken as an input and the external noise can be removed for better understanding of the input to analyse the speech and process the speech further to get desired results. After taking speech as input the speech recognition module helps to identify the spoken words and converts them into texts. Then further the words are processed in python backend where NLP techniques are being used. Firstly words are processed in tokenization then stem technique and bag_of_words technique .These process happens inside the neural network and we get the desired results being achieved by the assistant. API call is additionally processed because the assistant performs all the task that the user requests. For performing the actions the content extraction is being used . Also there is an additional feature that we added. The assistant can also recognize the emotion of a person by analysing the speech of a person that is being trained by the intents.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

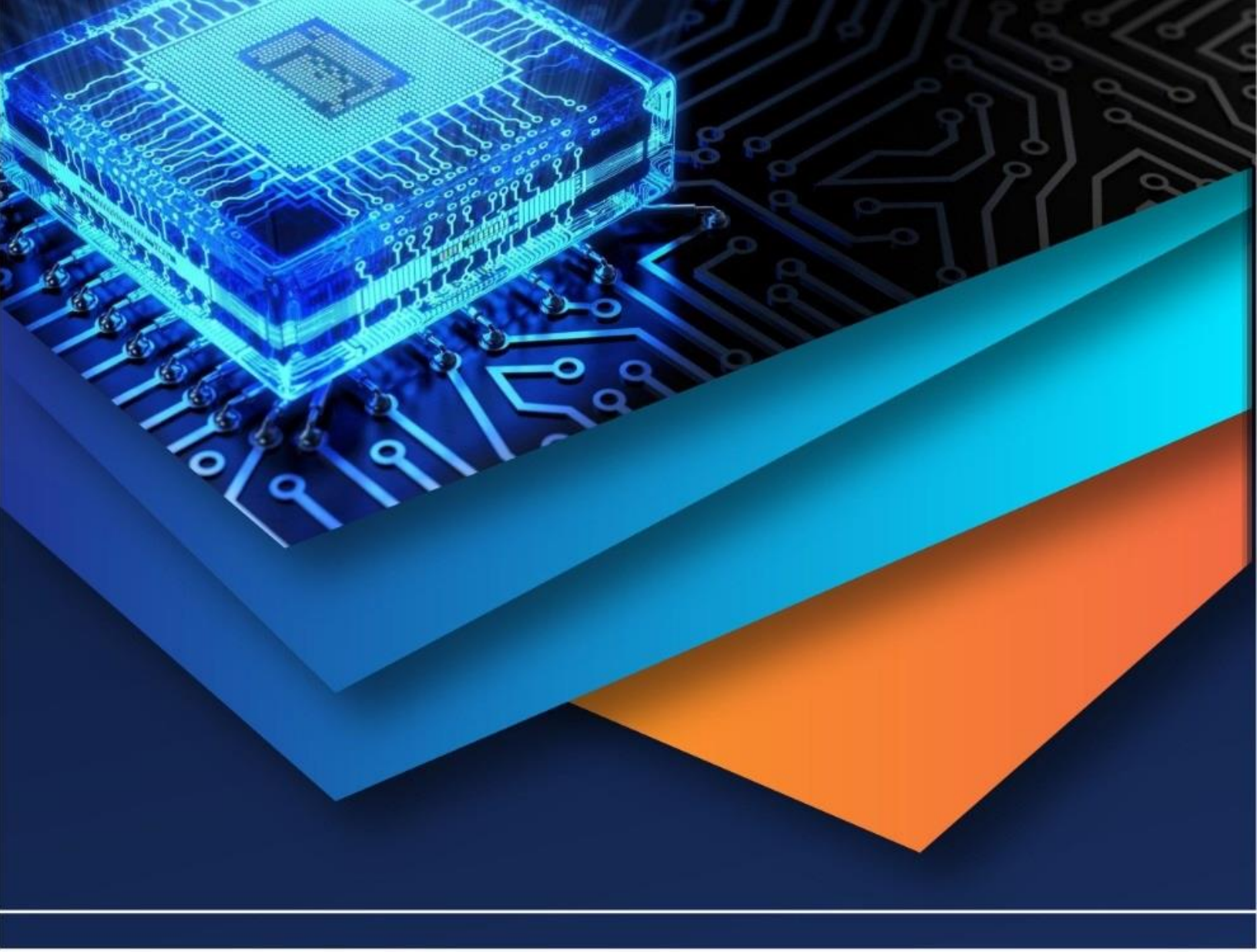
IV. CONCLUSION

The main objective of this paper is the AI Assistant can perform all the operations such as search engine tasks, setting remainders ,date, time, location, answering to the questions, playing youtube videos and also finds the emotion of a person by analysing the speech of a person. The scope of this paper is most of the people are busy in their work life they don't even find time to relax because of which their mental pressure can increase , so this AI Assistant can help them as this assistant is designed to talk with people and perform their desired tasks and also whenever people get bored or stressed they can communicate with this AI as this assistant can understand their emotion and talk with people in a friendly way.

REFERENCES

[1] https://www.researchgate.net/publication/320707512_Detection_and_Analysis_of_Human_Emotions_through_Voice_and_Speech_Pattern_Processing
[2] https://www.researchgate.net/publication/267229317_Human_Emotion_Recognition_System
[3] https://www.breitbart.com/tech/2017/05/02/new-artificial-intelligence-can-mimic-human-voices/

©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)