

**JOMO KENYATTA UNIVERSITY OF AGRICULTURE**

**AND TECHNOLOGY**

**DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING**

**BSc Electronic and Computer Engineering**

**FINAL YEAR PROJECT REPORT**

**PROJECT TITLE:**

**VOICE CONTROLLED ASSISTANT.**

**Submitted by:**

**MARK ODHIAMBO- ENE212-0154/2017**

**PROJECT SUPERVISOR**

**MR. ALOO**

*A Final Year Project Proposal submitted to the Department of Electrical and*

*Electronic Engineering in partial fulfillment of the requirements for the Award of a*

*Bachelor of Science Degree in Electronic Engineering.*

**NOVEMBER 2022**

# **DECLARATION**

This project proposal is my original work, except where due acknowledgement is made in the text, and to the best of my knowledge has not been previously submitted to Jomo Kenyatta University of Agriculture and Technology or any other institution for the award of a degree or diploma.

NAME.……………………MARK OKODE ODHIAMBO……………………….

SIGNATURE..…………………………………………DATE…21/11/22…………

**TITLE OF PROJECT: VOICE CONTROLLED ASSISTANT**

**SUPERVISOR CONFIRMATION:**

This project proposal has been submitted to the Department of Electrical and Electronic Engineering, Jomo Kenyatta University of Agriculture and Technology, with my approval as the University supervisor:

**NAME OF SUPERVISOR………**MR. ALOO**………………**

**SIGNATURE: ……………………………… DATE: ……………………………………….**

# ABSTRACT

Everyone who works has several tasks to do and sometimes can be overwhelming leading to burn out, stress and incomplete work. It is tedious to schedule, remember and manage a schedule especially when multiple tasks are involved it creates a tangled mess which can be frustrating. The aim of this project is to help solve this through a voice controlled assistant.

The voice controlled assistant is going to be used in scheduling, managing and setting reminders through voice commands. This will make work easier and ease the workload of an individual, through proper planning and management it aims to make people more productive and reduce stress.The project will also log the activities done by the voice assistant as a security feature through the enabling live monitoring of the voice assistants information logs.

This project has been implemented through the use of raspberry pi 4, Microphone and a speaker. The current existing technology in the market is mainly software programs that are not optimised for this specific niche. The project aims to be open source and offer the software program developed to all interested parties.

**TABLE OF CONTENTS.**

[**DECLARATION**](#_heading=h.30j0zll) **1**

[**ABSTRACT**](#_heading=h.1fob9te) **2**

[**1. INTRODUCTION.**](#_heading=h.3znysh7) **1**

[1.1 Background Information](#_heading=h.2et92p0) 1

[1.2 Problem statement](#_heading=h.tyjcwt) 1

[1.3 Justification.](#_heading=h.3dy6vkm) 2

[1.4 Objectives](#_heading=h.1t3h5sf) 2

[1.4.1 Main objectives](#_heading=h.4d34og8) 2

[1.4.2 Specific objectives](#_heading=h.2s8eyo1) 2

[**2. LITERATURE REVIEW**](#_heading=h.3rdcrjn) **3**

[2.1 Current systems.](#_heading=h.35nkun2) 3

[2.1.1 Technology used in current voice assistants.](#_heading=h.1y810tw) 5

[2.1.1.1 Speech to text.](#_heading=h.49x2ik5) 5

[2.1.1.2 Filtering out noise.](#_heading=h.19c6y18) 6

[2.1.1.3 Processing with neural networks](#_heading=h.1mrcu09) 7

[2.1.1.4 Understanding context with AI](#_heading=h.3l18frh) 8

[2.1.1.5 Evaluate response](#_heading=h.4k668n3) 8

[2.1.1.6 Text to speech](#_heading=h.1egqt2p) 8

[2.1.2 Shortcomings of current systems.](#_heading=h.4bvk7pj) 9

[2.1.2.1 Security and Privacy.](#_heading=h.2r0uhxc) 9

[2.3 Proposed System.](#_heading=h.kgcv8k) 9

[2.3.1 Improvements of the proposed system.](#_heading=h.43ky6rz) 10

[**2.5 Hardware components**](#_heading=h.fwqq8zq7cc4o) **10**

[2.5.1 Microprocessor.](#_heading=h.2nusc19) 10

[2.5.2 Raspberry pi.](#_heading=h.3mzq4wv) 10

[2.5.3 Speakers.](#_heading=h.40ew0vw) 11

[**3. METHODOLOGY**](#_heading=h.uhft60rz5if9) **13**

[3.1 Project Specification.](#_heading=h.1tuee74) 13

[3.2 Block diagram.](#_heading=h.279ka65) 13

[3.3 Flow chart.](#_heading=h.45jfvxd) 14

[3.4 Software development.](#_heading=h.4iylrwe) 16

[3.4.1 Application development](#_heading=h.2y3w247) 17

[3.4.2 API design](#_heading=h.3x8tuzt) 17

[3.4.3 Text processing program.](#_heading=h.2ce457m) 19

[**3.4.3.1 Voice Analysis.**](#_heading=h.holk1qsvuqzc) **19**

[**3.4.3.2 Transcription of the file**](#_heading=h.ars1pkffacgs) **20**

[**3.4.3.3 Text processing.**](#_heading=h.rb4xw3npwsud) **20**

[**3.5 Software tools and systems.**](#_heading=h.y3mxls7lm71i) **21**

[3.5.1 Python](#_heading=h.2w5ecyt) 21

[3.5.2 API](#_heading=h.3vac5uf) 21

[3.5.3 React](#_heading=h.39kk8xu) 22

[3.6 Hardware components](#_heading=h.rjefff) 22

[**4. RESULT AND ANALYSIS.**](#_heading=h.3bj1y38) **23**

[**4.1. HARDWARE DESIGN**](#_heading=h.b4jje0yl0nrj) **23**

[**4.2 SOFTWARE DESIGN ANALYSIS.**](#_heading=h.ygd0utbyoeer) **24**

[**4.2.1. AUDIO/ VOICE SAMPLE ANALYSIS.**](#_heading=h.qi60ucy5bkej) **24**

[**4.2.2 API RUNNING ON LOCALHOST.**](#_heading=h.o8njc3evbahx) **24**

[**4.2.3. WEB APPLICATION PLATFORM.**](#_heading=h.ykvb9cncjotn) **25**

[**5. CONCLUSION AND RECOMMENDATIONS.**](#_heading=h.u0yzfmuthzbo) **28**

[**5.1 CONCLUSION**](#_heading=h.y1f3qqh5be8c) **28**

[**5.2 RECOMMENDATION.**](#_heading=h.blbszhw3qu1d) **28**

[**6. BUDGET**](#_heading=h.4sernu590mj5) **29**

[**TIME-PLAN**](#_heading=h.2pta16n) **30**

[**REFERENCES**](#_heading=h.14ykbeg) **31**

**LIST OF FIGURES**

Figure 2.1:Alexa 4

Figure 2.2:Analogue to digital conversion 7

Figure 2.3:Digital to analogue conversion 8

Figure 2.4:Raspberry pi 12

Figure 2.5:ReSpeaker 2-Mic PiHAT 12

Figure 3.1:Block diagram 13

Figure 3.2:Flow chart 14

Figure 3.3:Web platform home page 24

Figure 3.4:Todo api sample code 25

Figure 3.5:Todo Endpoint 25

Figure 3.6:Audio analysis code 26

Figure 3.7:Transcription of the audio file 27

Figure 3.8:Text processing system 28

Figure 4.1:The hardware system design implemented 30

Figure 4.2:Plot of a sample wav file 31

Figure 4.3:API running on the console 32

Figure 4.4:Calendar page 33

Figure 4.5:Activity Log page 33

Figure 4.6:Todo list page 34

**LIST OF ABBREVIATIONS AND ACRONYMS**

API Application Programmable Interface

**CHAPTER ONE**

# 1. INTRODUCTION.

## 1.1 Background Information

Task and schedule management has been a concern in numerous industries for years, including finance, health care, administration, and retail. In the current technological age, it is not only necessary to manage the existing tasks, but it is also important to boost productivity. This problem is faced with many people. Therefore, this project suggests that using a Voice controlled assistant can boost productivity in the workforce.

Scheduling tasks is one of the most important activities in today's fast based work environment. Productivity can be boosted with a well designed voice controlled assistant. The implementation of a voice controlled assistant will help a person to plan, schedule and manage tasks.

## 1.2 Problem statement

People who work will always schedule, manage and plan their tasks. At some point, everyone will need to boot their productivity. A properly made schedule is critical and a poorly made one will end up causing harm, confusion and stress. The majority of either follow a timetable which is static and not ideal for a flexible work schedule. People require a platform that will enable them to be able to be more dynamic and allow them to be flexible in their planning. The voice controlled assistant aims to achieve this.

In reality, most people rarely understand the benefits of time management and proper planning and scheduling. We rely on a time table and setting appointments which can lead to conflict of different tasks at the same time and leading to misunderstandings , which is a sort of inconvenience. As a result, a lot of time is wasted that could be spent on more important activities. Implementing a voice controlled assistant would aid in this.

The proposed voice controlled assistant aims to improve creativity and help with time management. They won't have to spend as much time on setting appointments through different software or worse an analogue system of writing things down which can be misplaced or lost, when they could be doing something more productive.

## 1.3 Justification.

The project is significant because it has assisted to schedule and set tasks through this has increased productivity with the use of a voice controlled assistant. Productivity is important because it will help one accomplish more tasks in less time and more efficiently. This project has been measured on its ability to take in tasks and the accuracy in which this happens.

The project has introduced a two access point system to the voice assistant, increased security of access control systems through the use of monitoring and has been able to be able to interface with an application programmable interface(API) allowing it to be able to post information that could be accessed at a different endpoint(web application).

## 1.4 Objectives

### 1.4.1 Main objectives

To design, implement and test a voice controlled assistant, with a real-time update of setting tasks and reminders to boost productivity.

### 1.4.2 Specific objectives

1. To design and deploy an application display for task scheduling and time management.
2. To design and test an electronic system for voice detection, synthesize and processing so as to be able to query tasks.
3. To implement and prototype an electronic system for voice output.

**CHAPTER TWO**

# 2. LITERATURE REVIEW

Voice assistants are software agents that can interpret human speech and respond via synthesized voices. Apple’s Siri, Amazon’s Alexa, Microsoft’s Cortana, and Google’s Assistant are the most popular voice assistants and are embedded in smartphones or dedicated home speakers. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands.

The personal assistant is developed on a microprocessor based platform. Recent advances in natural language processing, also known as computational linguistics, has allowed voice assistants to create meaningful responses quickly. Hirschberg and Manning credit these recent improvements in natural language processing to four things.A vast increase in computing power, the availability of very large amounts of linguistic data, the development of highly successful machine learning methods, and a much richer understanding of the structure of human language and its deployment in social contexts.[1]

In this project, we develop a voice-controlled assistant. The human voice commands are given to the assistant remotely, by using an audio system. The assistant can manage several tasks such as setting appointments and a schedule. The voice commands are processed in real-time, using an online cloud server for example google. The speech signal commands converted to text form are communicated to an API the tasks are later logged and displayed on the app.

## 2.1 Current systems.

Apple’s Siri, Microsoft’s Cortana, Amazon’s Alexa, and Google’s Assistant are all software agents that run on purpose-built speaker devices or smartphones. The software constantly listens for a keyword to wake it up. Once it hears that keyword, it records the user’s voice and sends it to a specialized server, which processes and interprets it as a command. Depending on the command, the server will supply the voice assistant with appropriate information to be read back to the user, play the media requested by the user, or complete tasks with various connected services and devices. The number of services that support voice commands is growing rapidly, and Internet-of-Things device manufacturers are also building voice control into their products.

[2]Apple’s Siri assistant has been around the longest, released as a standalone app in 2010 and bundled into iOS in 2011. Microsoft followed shortly thereafter with Cortana in 2013. Amazon launched Alexa with its Echo-connected home speaker in 2014, and Google’s Assistant was announced in 2016 along with its home speaker and is also embedded in the Google app for Android-based smartphones. Each assistant has its own unique features, but the core functions are the same. Voice assistants differ from earlier voice-activated technologies in that they can respond to a much larger number of commands and questions. This is because they are always connected to the Internet; each interaction is sent back to a central computing system that analyzes the user’s voice commands and provides the assistant with the proper response. Earlier voice-activated devices relied on a smaller set of “built-in” commands and responses. Recent advances in natural language processing, also known as computational linguistics, has allowed voice assistants to create meaningful responses quickly. An example of a voice controlled assistant as shown in the figure 2.1 below.



*Figure 2.1: Alexa*

There is also the famous alexa that is capable of doing many tasks on top of searching information from the web. It is capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audiobooks, and providing weather, traffic, sports, and other real-time information, such as news. Alexa can also control several smart devices using itself as a home automation system.

### 2.1.1 Technology used in current voice assistants.

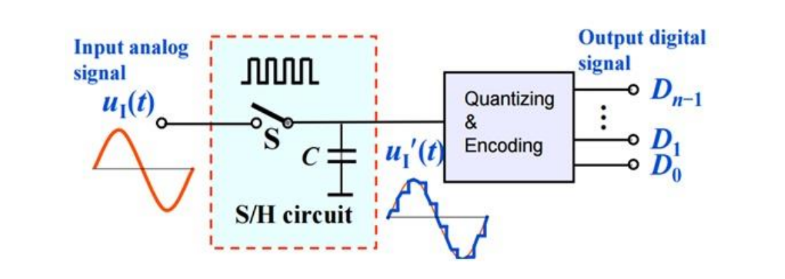
The voice assistant is based on understanding the human language and interpreting the same to offer appropriate results. AI programming perfects its algorithms to constantly provide the best rational answer. A mixture of AI and automation helps develop speech systems. For example when two people communicate, there is encoding and decoding of the message; voice assistants work similarly.

Several technologies are used :

* Speech to text
* Filtering out noise
* Processing with neural networks
* Understanding context with AI
* Evaluate response. (Text generation)
* Text to speech

#### 2.1.1.1 Speech to text.

When discussing [3]digital communication principles there is analog to digital conversion. This is the process in which an analogue signal is converted into a digital signal. By converting from the analog world to the digital world, we can begin to use electronics to interface to the analog world around us. As shown in the figure 2.2.



*Figure 2.2: Analogue to digital conversion.*

There are two steps for the analog to digital conversion:

1. Sampling and Holding.

An analog signal continuously changes with time, in order to measure the signal we have to keep it steady for a short duration so that it can be sampled. We can hold the signal for a specific duration and then digitize the signal and sample the value. This is done by a sample and hold circuit. An important consideration is the Nyquist theorem.According to the Nyquist theorem, the sampling rate must be at least 2 times the highest frequency contained in the signal. It is also known as the minimum sampling rate and given by: Show in the Equation(2.1)

(2.1)

2. Quantizing and Encoding.

In order to complete analog to digital conversion, each sample value is mapped to a discrete level (represented by a sequence of bits) in a process called quantization. Quantization is a method of representing the sampled values of the amplitude by a finite set of levels, it’s like rounding off. It converts a continuous-amplitude sample to a discrete-amplitude sample. Quantization is used to limit the amplitude of these samples.

formula for height of the L levels: Shown in the example(2.2)

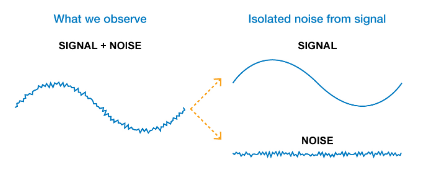
………………………..(2.2)

L in the equation is the quantization level.

One of the applications of analogue to digital conversion is speech to text. The voice which is analogue information is converted to digital information then analyzed into text information.

#### 2.1.1.2 Filtering out noise.

There are chances that in addition to the words spoken by the user, some ambient sound is picked up by the system. Say connecting with a call center while on the road can increase the chances of recording the surrounding disturbances like horns or announcements going on in the message. This is noise.[9] Noise is an unwanted signal, which interferes with the original message signal and corrupts the parameters of the message signal. This alteration in the communication process, leads to the message getting altered. The unwanted signal superimposes itself on the message signal. Noise is illustrated in the figure 2.3.



*Figure 2.3: Noise signal*

[10]Noise is therefore removed through the use of analogue filters and in the case of voice assistants there is the use of machine learning in the denoising of the signal

#### 2.1.1.3 Processing with neural networks

[8]Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of [machine learning](https://www.ibm.com/cloud/learn/machine-learning) and are at the heart of [deep learning](https://www.ibm.com/cloud/learn/deep-learning) algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.

The voice assistant is based on the neural networks. Reading and analyzing every single letter of the message, the AI tries to analyze the sentence’s meaning and match it with the best possible outcomes.

Neural networks require a large corpus of data so that they are able to make inferences from and give the expected or the desired output.

#### 2.1.1.4 Understanding context with AI

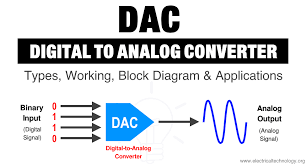
The voice AI is now ready to act. Using syntactic and semantic techniques for analyzing text, the AI gets a deeper understanding of the [context](https://verloop.io/blog/user-intent-in-ai/) under consideration. Here, the syntactic analysis further breaks down the natural language for the grammatical rules. Moreover, semantic analysis is based on understanding the meaning of the sentences and words.

#### 2.1.1.5 Evaluate response

The voice assistant reaches a specific range of conclusions by carefully examining the question of the user. The algorithm further analyses the most potential solution and filters the responses to find the perfect match to the query. Therefore finally generating a text to respond to the user.

#### 2.1.1.6 Text to speech

[7]The final step is the text to speech . This is the process of converting digital information into analog information which is speech in the case of the voice assistant. In current voice assistants systems the text to speech systems have a high level human mimicry of the human voice, for example,Siri. The system is shown in the figure 2.4.



*Figure 2.4: Digital to analog conversion*

### 2.1.2 Shortcomings of current systems.

#### 2.1.2.1 Security and Privacy.

While voice assistants have interesting and useful features, they also pose several unique problems. One of the main issues with these voice-activated devices is security. Anyone with access to a voice-activated device can ask it questions, gather information about the accounts and services associated with the device, and ask it to perform tasks. This poses a major security risk because these devices will read out calendar contents, emails, and other highly personal information. In one reported case, a man discovered that the iPad in his living room would unlock the front door for anyone who stood outside and asked Siri to let them in.[4]

Amazon’s Alexa is just as prone to these security issues, and Amazon is working to deploy a similar voice printing system. Alexa has the added issue of being built into Amazon’s store interface. By default, anyone with voice access to the device can order items using the owner’s Amazon account. There are options to set a voice passcode to confirm purchases, and all goods will ship to the owner’s address on file, but there is still potential for malicious users to purchase goods on the owner’s account. Household members could make unauthorized purchases as well, like the six-year-old who ordered herself a dollhouse and four pounds of sugar cookies via Alexa.[5]

Voice assistants are also vulnerable to several other attacks. Researchers have recently proven that voice assistants will respond to inaudible commands delivered at ultrasonic frequencies.[6]This would allow an attacker to approach a victim, play the ultrasonic command, and the victim’s device would respond. There is also the possibility this type of attack could be embedded in broadcast media.This is known as a dolphin attack. DolphinAttack, that modulates voice commands on ultrasonic carriers (e.g., f > 20 kHz) to achieve inaudibility.

## 2.3 Proposed System.

The voice controlled assistant that is to be developed in this project is a voice controlled assistant with the ability of transcribing information and querying tasks for the tasks schedule and project management.

The voice controlled assistant should be able to display the tasks set on an application so that the user can be able to have a visual display of the tasks on the application; this is a unique feature of the project.

### 2.3.1 Improvements of the proposed system.

The improvement of the voice assistant are:

* The voice controlled assistant is able to interface with an API enabling the device to have two access points therefore increasing accessibility.
* Through the two access points the admin user will be able to see the log of the activities in which the voice controlled assistant is undertaking therefore using this as a security feature.
* The ability of the voice controlled assistant to interface with an API therefore enables it to be able to develop a task management application as a web application for the purpose of time management.

## 2.5 Hardware components

### 2.5.1 Microprocessor.

A microprocessor is a computer processor where the data processing logic and control is included on a single integrated circuit, or a small number of integrated circuits. The microprocessor contains the arithmetic, logic, and control circuitry required to perform the functions of a computer's central processing unit.

### 2.5.2 Raspberry pi.

Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation in association with Broadcom. The Raspberry Pi project originally leaned towards the promotion of teaching basic computer science in schools and in developing countries. Currently raspberry pi is used in small prototyping projects due its advantages and cost.

Raspberry pi was used as the microprocessor because several important features were used for the input. This is for example its ability to have a higher computation power that would be useful to the process of the voice command and factoring that there will be also other components such as the speaker and the microphone. The microprocessor is shown in figure 2.4.

****

*Figure 2.4: Raspberry pi*

### 2.5.3 Speakers.

This is a device that is responsible for the conversion of the electrical energy into a sound energy. The speaker for this system is used so that the computer will be able to communicate information and queries to the user.

**2.5.4 ReSpeaker 2-Mics Pi HAT**

ReSpeaker 2-Mics Pi HAT is a dual-microphone expansion board for Raspberry Pi designed for AI and voice applications. This means that you can build a more powerful and flexible voice product that integrates Amazon Alexa Voice Service, Google Assistant, and so on. The board is developed based on WM8960, a low power stereo codec. There are 2 microphones on both sides of the board for collecting sounds and it also provides 3 APA102 RGB LEDs, 1 User Button and 2 on-board Grove interfaces for expanding your applications. What is more, 3.5mm Audio Jack or JST 2.0 Speaker Out are both available for audio output. The component is displayed in the figure 2.5.



*Figure 2.5: ReSpeaker 2-Mics Pi HAT*

# 

# 

# 3. METHODOLOGY

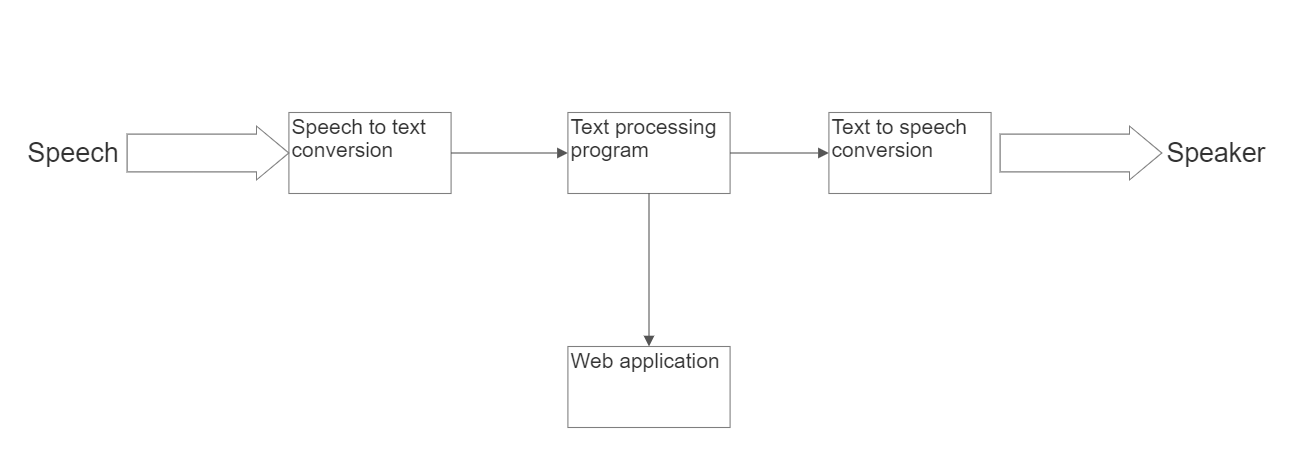
## 3.1 Project Specification.

The systems specification are:

* A raspberry pi 4 having a 1.5 GHz 64-bit quad core [ARM Cortex-A72](https://en.wikipedia.org/wiki/ARM_Cortex-A72) processor, on-board 802.11ac [Wi-Fi](https://en.wikipedia.org/wiki/Wi-Fi), [Bluetooth 5](https://en.wikipedia.org/wiki/Bluetooth_5), full [gigabit Ethernet](https://en.wikipedia.org/wiki/Gigabit_Ethernet) (throughput not limited), two [USB 2.0](https://en.wikipedia.org/wiki/USB_2.0) ports, two [USB 3.0](https://en.wikipedia.org/wiki/USB_3.0) ports, 1–8 GB of RAM, and dual-monitor support via a pair of micro HDMI ([HDMI Type D](https://en.wikipedia.org/wiki/Mini-HDMI)) ports for up to [4K resolution](https://en.wikipedia.org/wiki/4K_resolution).
* A microphone to take in the voice from the user and process it to get text information.
* Speaker to convert the Text to voice output.
* An API that POST from the voice command and GET from the application end and the voice program.
* An application for task management and time planning.
* A test processing and analysis of the user voice input.

## 3.2 Block diagram.

The project has taken in the audio input(speech) process it then retrieve information if there is a tasks it has been sent the API to connected to the web application and posted it then based on the context the text processing program has evaluate and give and output which will then be converted to the speech and released through the speaker. The block diagram is shown in figure 3.1.

****

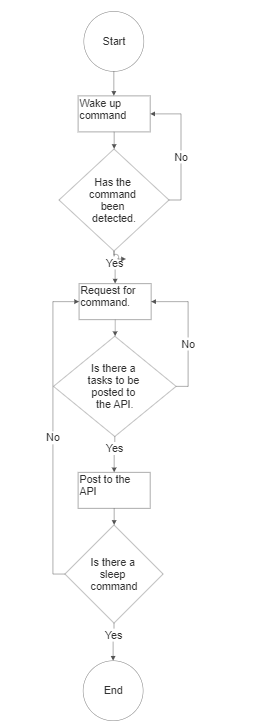
*Figure 3.1 Block diagram*

## 

## 3.3 Flow chart.

The flow chart in the figure 3.2. This illustrates the information flow of the project. The steps of follows:

* The project has checked for a wake up command if no the project returns awaiting for a wake up command and if present it proceeds to request for the command from the user.
* The command is processed and if there is a task present it proceeds and sends it to API if not it requests for the command again.
* The project proceeds to query for the sleep command if present the project ends and if there is no sleep command it remains active.

****

*Figure 3.2:Flow chart*

## 3.4 Software development.

Through the use of the flow chart it can be seen the flow of information of how the project has run. The software will consist of the three main parts.

* Application development
* API design.
* Text processing program

### **3.4.1 Application development**

This is tied to the objective number one; to design and implement an application for task schedule management with the goal of time planning to increase productivity. The application should run on the web and connect to the API.



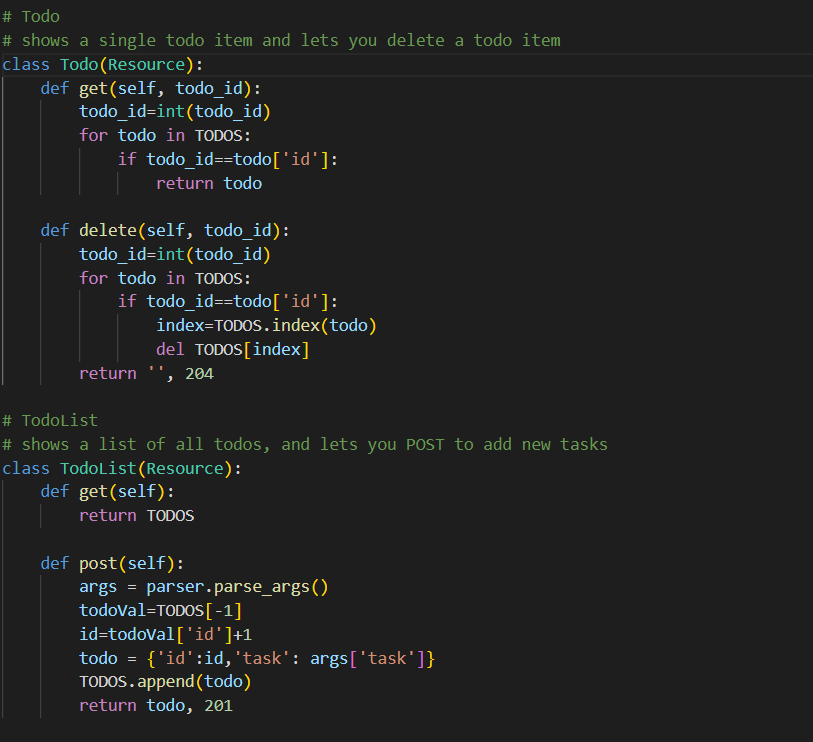
*Figure 3.3 Web platform home page*

The website has home tabs for the various objectives it had: the home, todo, remainder/calendar and finally the activity log.

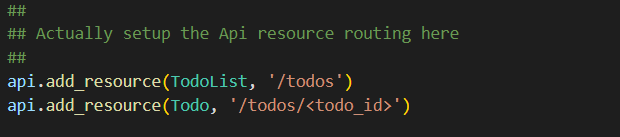
### 3.4.2 API design

The API has acted as a bridge between the hardware and the application. From the hardware the user has been able to post data and from the application side the user will get the information. The api was coded in flask restful and was routed into the different endpoint.

Below is an example of the todo api that was used in figure 3.4 and the endpoint in figure 3.5.



*Figure 3.4 Todo api sample code.*

**

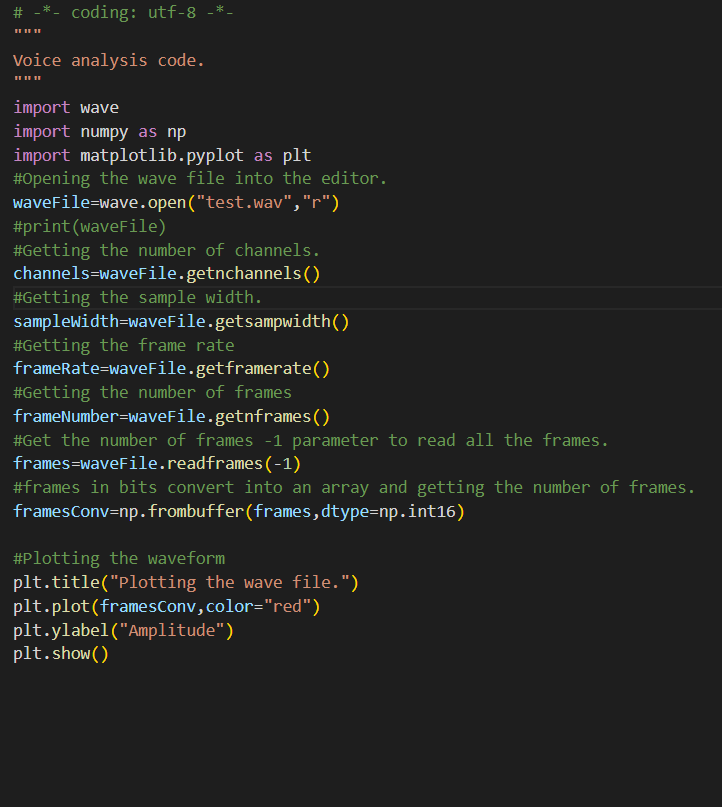
*Figure 3.5 Todo endpoints*

### 3.4.3 Text processing program.

This has taken the speech text, analyzed and retrieved the information and enabled the machine to understand the path in which it should respond.

#### 3.4.3.1 **Voice Analysis.**

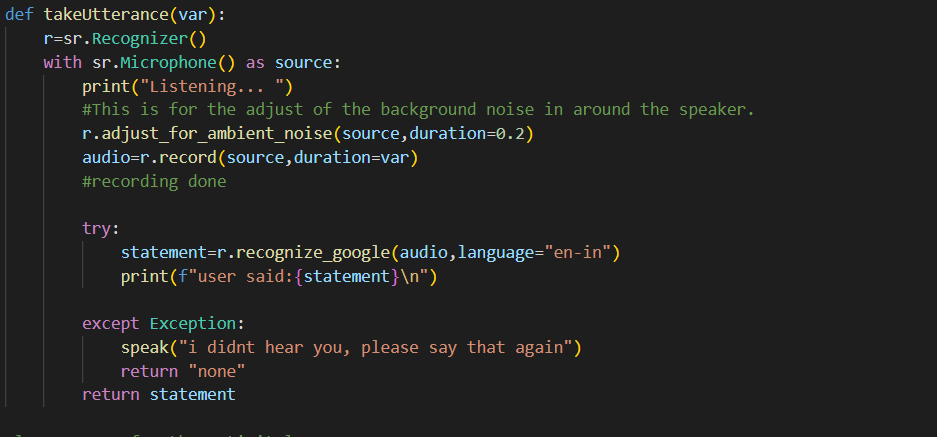
This is the voice analysis of the sound file that is the system. This is the audio file that will later be transcribed by the program.



*Figure 3.6. Audio analysis code.*

#### 3.4.3.2 Transcription of the file

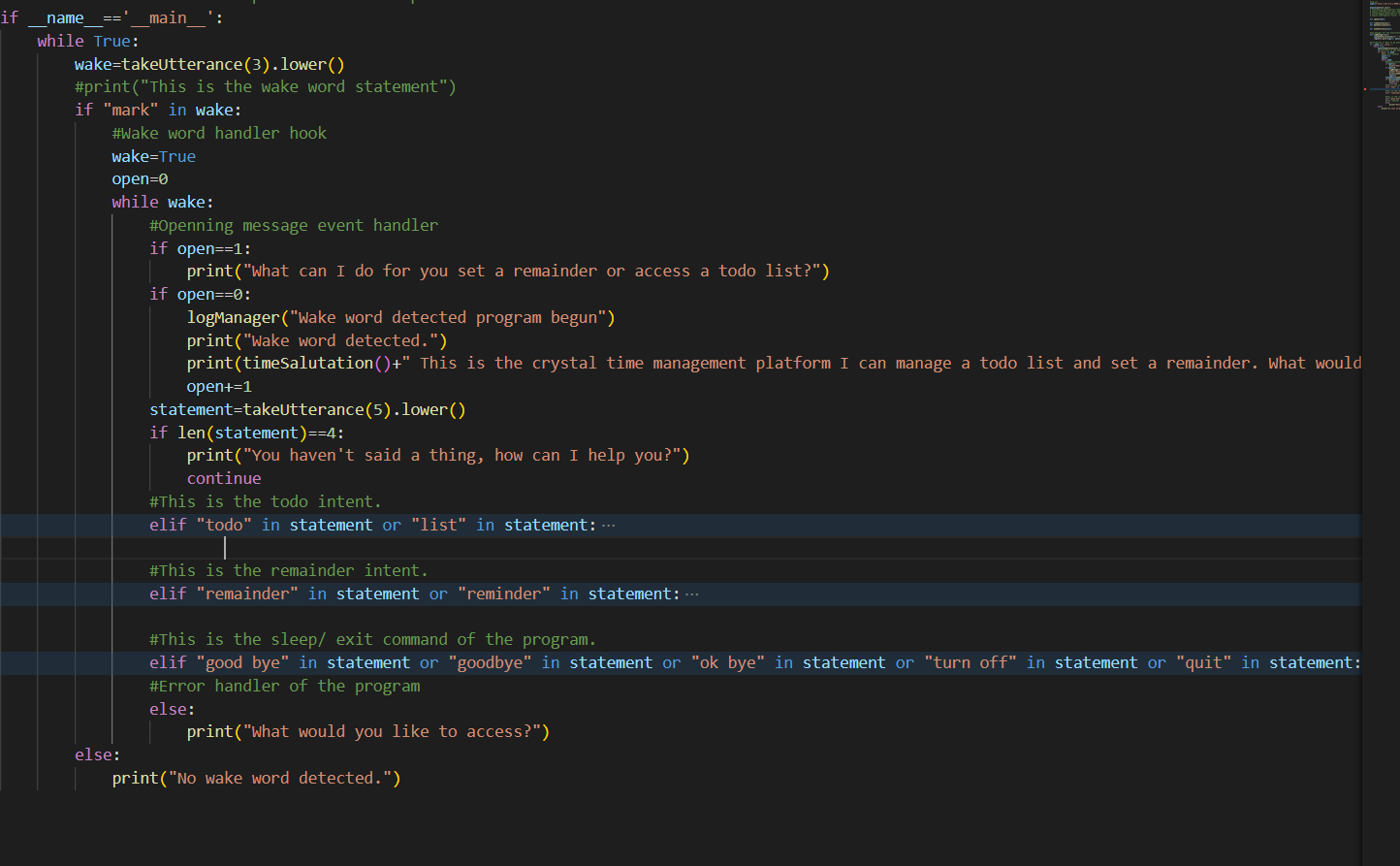
Transcription is the process of converting an audio file into a text file. This was done through the use of the speech\_recogniver library with the use of a google api.



*Figure 3.7 Transcription of the audio file.*

#### 3.4.3.3 Text processing.

The system used in the program for text processing was a keyword system in the program below it can be noted that each time a specific keyword was spoken the program run a specific block of code.



*Figure 3.8 Text processing system.*

## 3.5 Software tools and systems.

This is the review of the software selection of the project and the reason for selecting the software.

### 3.5.1 Python

In the project the microprocessor will run on the platform through the use of the Python language. Python is a high-level, interpreted, general-purpose programming language. We chose python because it has a shallow learning curve, its syntax and semantics are transparent, and it has good string-handling functionality.[7] Python comes with a standard library, including components for graphical programming, numerical processing and web connectivity.

### 3.5.2 API

API is the acronym for Application Programming Interface, which is a software intermediary that allows two applications to talk to each other. Each time you use an app like Facebook, send an instant message, or check the weather on your phone, you’re using an API.

When you use an application on your mobile phone, the application connects to the Internet and sends data to a server. The server then retrieves that data, interprets it, performs the necessary actions and sends it back to your phone. The application then interprets that data and presents you with the information you wanted in a readable way. This is what an API is - all of this happens via API.

### 3.5.3 React

The react will be used for the front-end and will enable the information to be updated in real-time onto the person’s device as the voice commands are issued. This is the open source javascript library that is used for building the user interface.

## 3.6 Hardware components

Considering the necessary requirements, the chosen solution to build the voice controlled assistant will be based on the following components.

* Speaker
* ReSpeaker 2-Mics Pi HAT
* Raspberry pi 4

**CHAPTER FOUR**

# 4. RESULT AND ANALYSIS.

## 4.1. HARDWARE DESIGN

This project is a voice controlled assistant that has been optimized for time management. Its hardware design is as shown below.



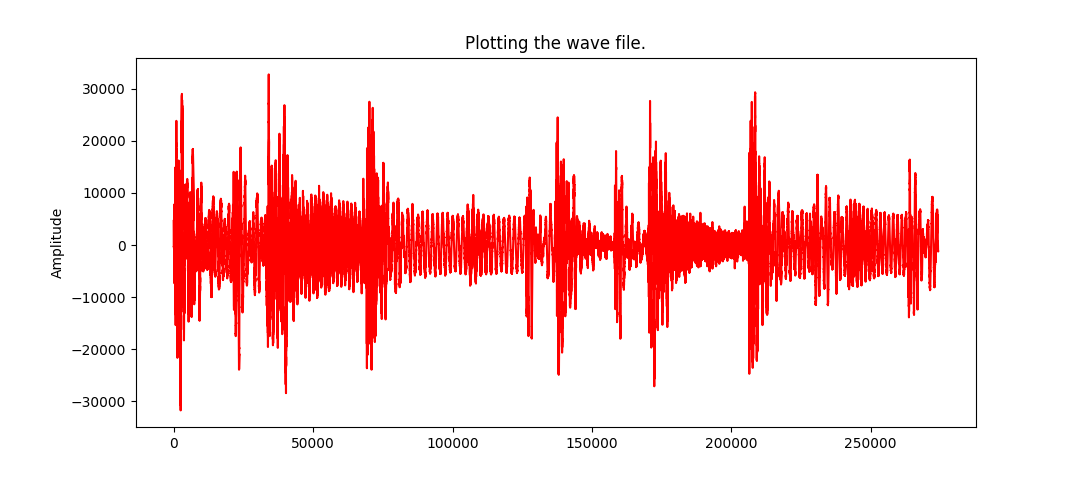
*Figure 4.1 The hardware system design implemented.*

## 4.2 SOFTWARE DESIGN ANALYSIS.

The Raspberry pi was set up and the sound card was set up in the raspberry pi and the components were then tested.

## 4.2.1. AUDIO/ VOICE SAMPLE ANALYSIS.

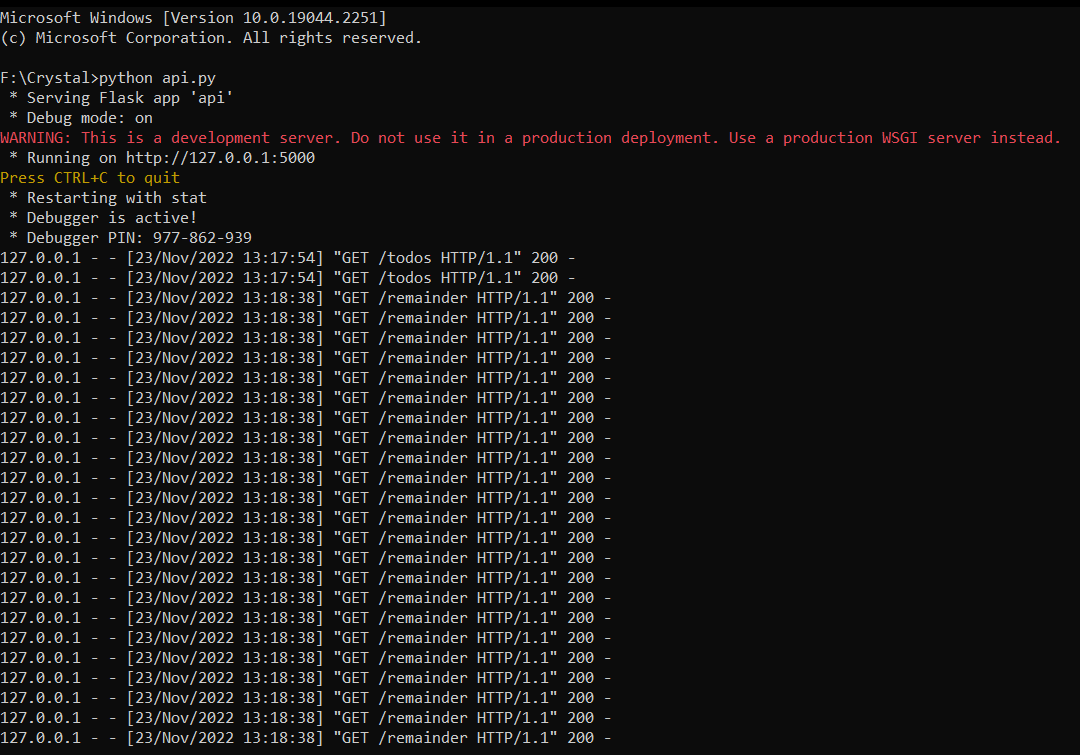
This is the analysis of the audio sample. A plot of a sample wav file is shown in figure 4.2



*Figure 4.2. Plot of a sample wav file.*

### 4.2.2 API RUNNING ON LOCALHOST.

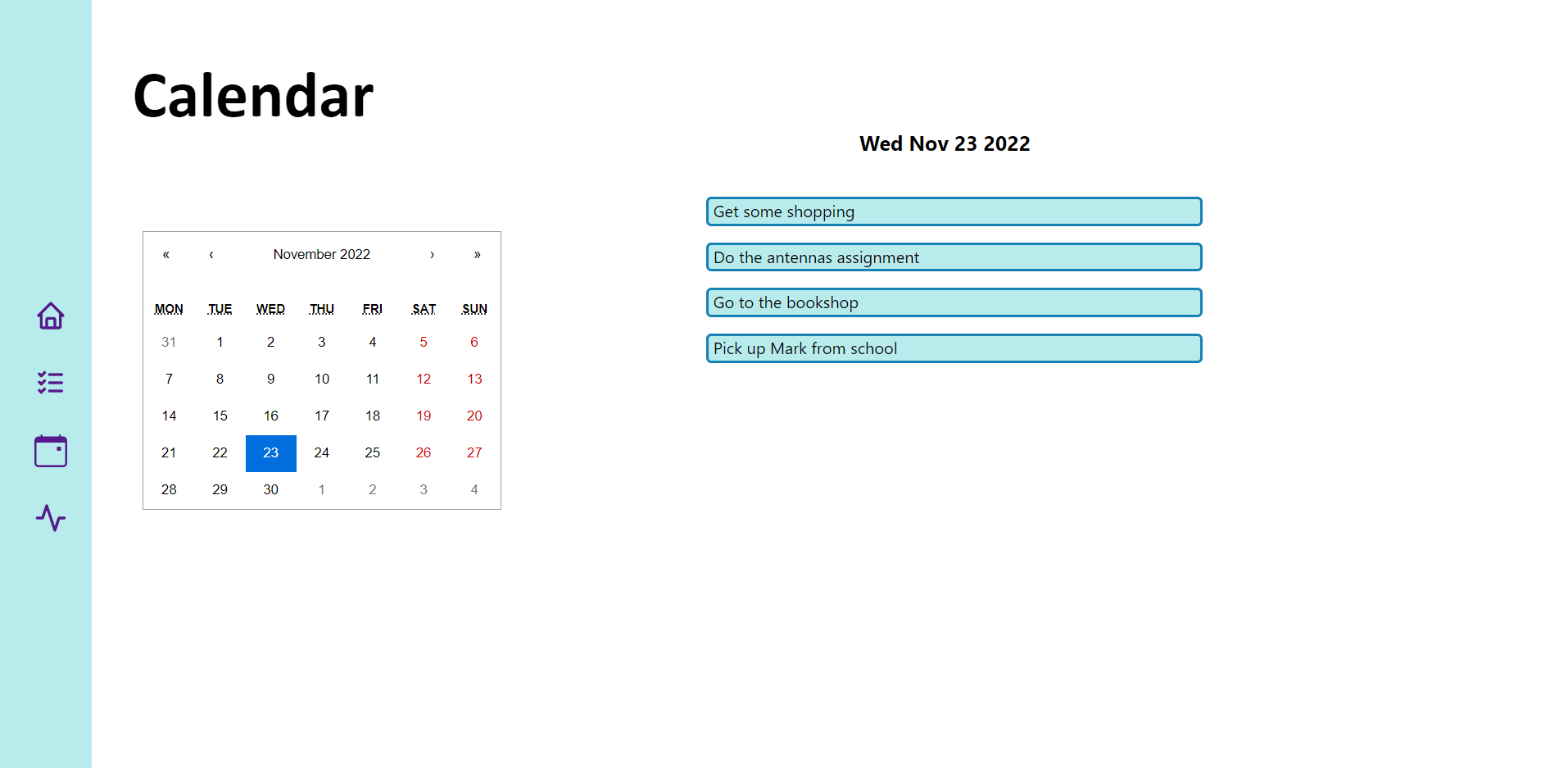
The api was running on the localhost to allow for ease with debugging and access for prototyping. The sample of the running console is shown in the figure 4.3.



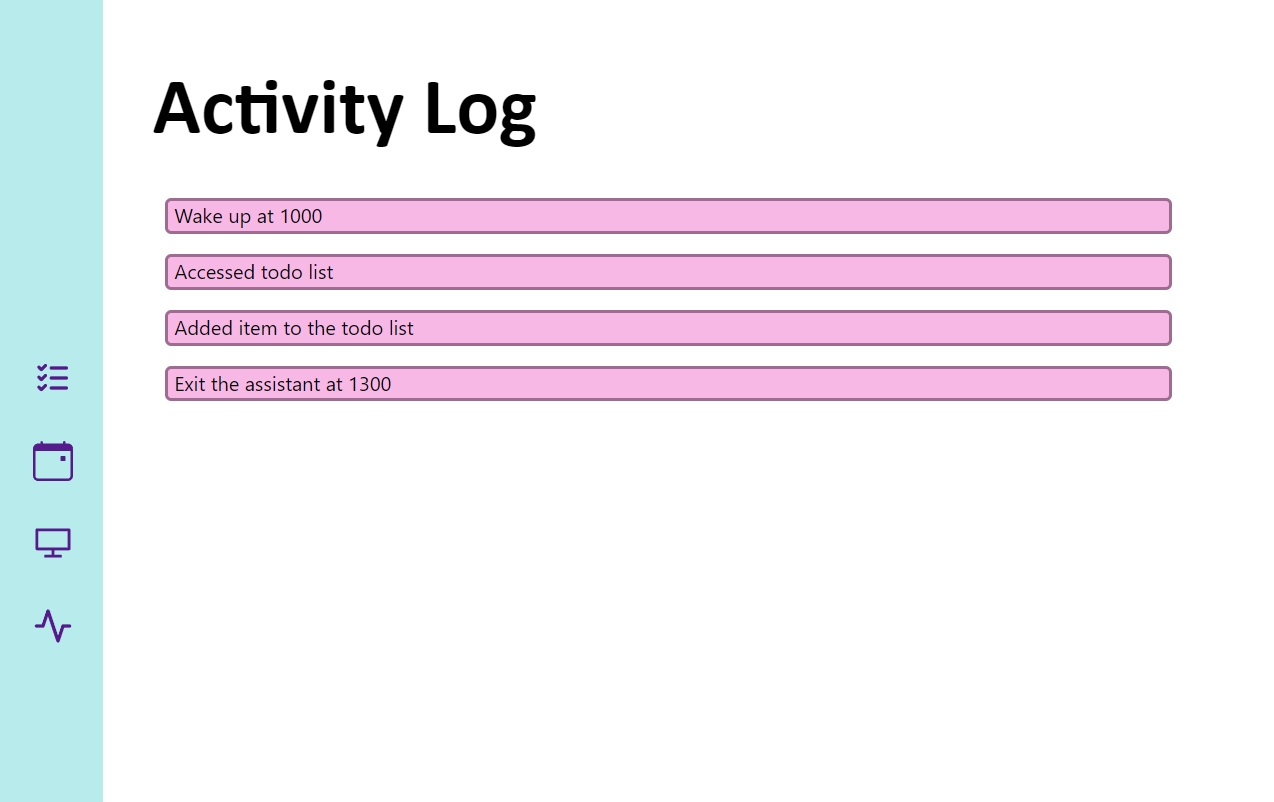
*Figure 4.3. API running on the console.*

### 4.2.3. WEB APPLICATION PLATFORM.

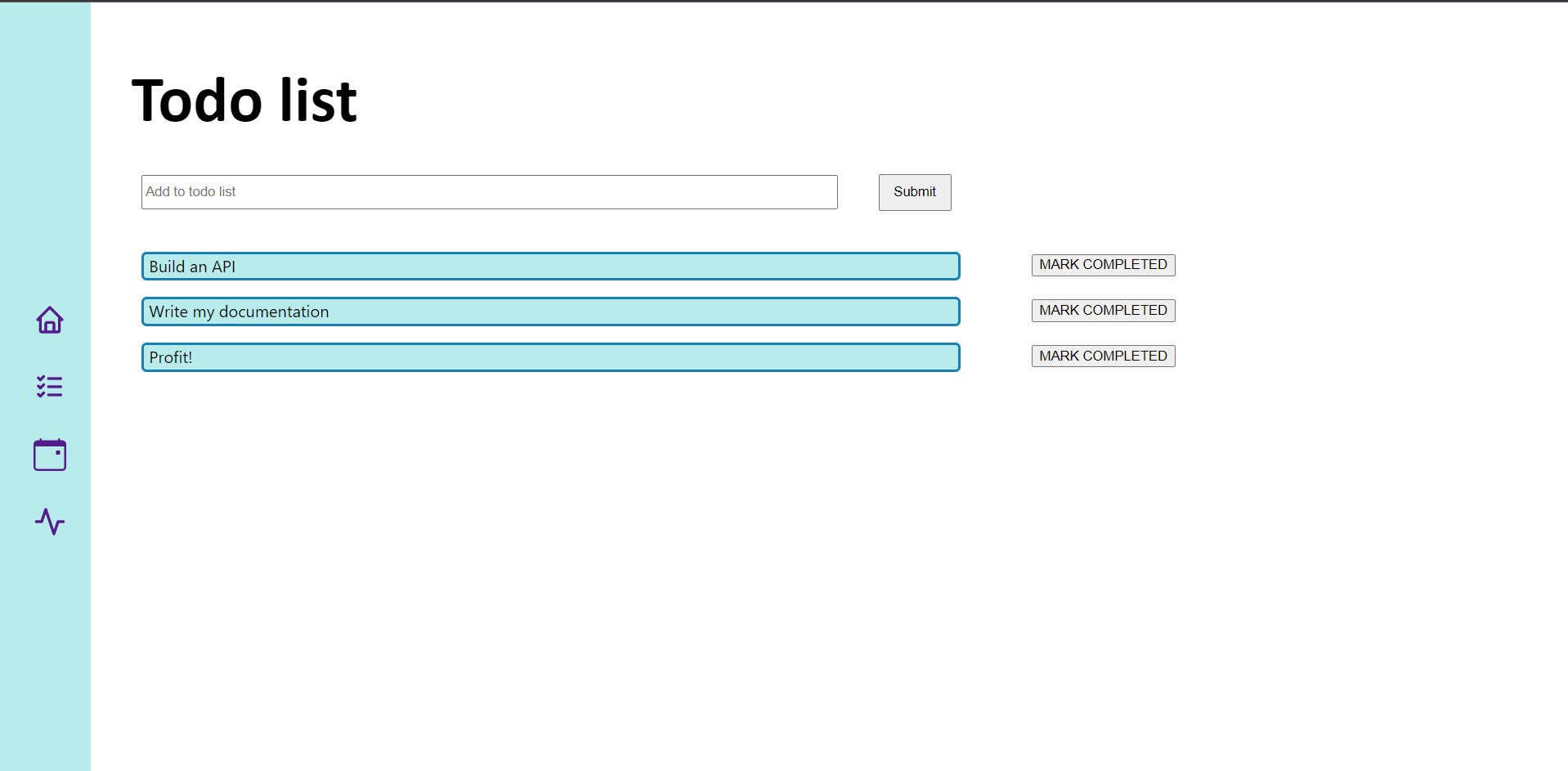
The web platform was designed and implemented below is a sample of how the result of the webpage. Figure 4.4 shows the calendar page , figure 4.5 shows the activity log and then figure 4.6 shows the todo list page.



*Figure 4.4 Calendar page.*



*Figure 4.5 Activity log page.*

**

*Figure 4.6 Todo list page*

# 

# 

**CHAPTER FIVE**

# 5. CONCLUSION AND RECOMMENDATIONS.

## 5.1 CONCLUSION

In this project, a voice controlled assistant has been designed and prototyped. From the design steps taken the user has been able to take voice/audio information. The project utilized a microphone and speaker to enable communication.

The project was able to post information to the api and the website was able to retrieve the information. The system was a 2 point system since the website had several functionalities to allow for the user to manage the tasks from the website and voice end.

Finally the project objectives were met as per the indicated methodology and hence analysis for the final designed voice controlled assistant. Implementation of it would help would optimise time use.

## 5.2 RECOMMENDATION.

Given the scope of this project’s objectives, other opportunities that arise to enhance the design of the voice controlled assistant include;

* Algorithm for transcription of audio file
* Microphone of increased range.
* Text to speech algorithm to allow for the machine to sound more human.

# 6. BUDGET

| Item | Description | Quantity | Rate | Amount |
| --- | --- | --- | --- | --- |
| 1 | Raspberry pi 4 | 1 | 13000 | 13000 |
| 2 | Speaker | 1 | 1200 | 1200 |
| 3 | Microphone | 1 | 4000 | 4000 |
| 4 | Wires | 1 | 800 | 800 |
| TOTAL | | | | 10600 |

# TIME-PLAN

| **ACTIVITIES** | **MAY 2022** | **JUNE2022** | **JULY 2022** | **AUG 2022** | **SEP** **2022** | **OCT 2022** | **NOV 2022** | **DEC 2022** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Documentation** |  |  |  |  |  |  |  |  |
| **Proposal Writing** |  |  |  |  |  |  |  |  |
| **Literature Review** |  |  |  |  |  |  |  |  |
| **Proposal Presentation** |  |  |  |  |  |  |  |  |
| **Design and coding** |  |  |  |  |  |  |  |  |
| **Hardware configuration, testing and adjustment** |  |  |  |  |  |  |  |  |
| **Final Report writing** |  |  |  |  |  |  |  |  |
| **Final Presentation** |  |  |  |  |  |  |  |  |

# REFERENCES

[1] Hirschberg, Julia, and Christopher D. Manning. “Advances in Natural Language Processing.” *Science* 349 (6245): 261–266. doi:10.1126/science.aaa8685.

[2]Matthew B. “Hoy Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants, Medical Reference Services Quarterly”(2018) , 37:1, 81-88.

[3]Tuan Do-Hong. “Principles of Digital Communications “(2007)

[4] Tilley, Aaron. “How A Few Words to Apple’s Siri Unlocked a Man’s Front Door.” *Forbes* (September 21, 2016). <https://www.forbes.com/sites/aarontilley/2016/09/21/apple-homekit-siri-security/>.

[5]Liptak, Andrew. “Amazon’s Alexa Started Ordering People Dollhouses after Hearing Its Name on TV.” *The Verge* (January 7, 2017). <https://www.theverge.com/2017/1/7/14200210/amazon-alexa-tech-news-anchor-order-dollhouse>.

[6] Zhang, Guoming, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. “DolphinAttack: Inaudible Voice Commands.” In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, New York, NY, 103–117. doi:10.1145/3133956.3134052.

[7]Steven Bird, Ewan Klein, Edward Loper. “Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit,”2009.

[8]Ethem Alpaydin. “Introduction to Machine Learning,” MIT Press, 2020.

[9]Hernando Lautaro Fernandez-Canque. “Analog Electronics Applications Fundamentals of Design and Analysis, ” Boca Raton, 2016.

[10]Hercules G. Dimopoulos.”Analog Electronic Filters: Theory, Design and Synthesis”,Springer Science & Business Media, 2011.