**AI ASSISTANT**

**PROJECT SYNOPSIS**

OF MAJOR PROJECT

**BACHELOR OF TECHNOLOGY**

## Branch CSE

SUBMITTED BY:

RISHIKA GUPTA 2000290100117

SAJAL GUPTA 2000290100126

SAKSHAM PANDIT 2000290100128

August 2022



## **KIET Group of Institutions, Delhi-NCR,**

## **Ghaziabad (UP)**

## **Department of Computer Science and Engineering**

**Table of contents**

|  |  |
| --- | --- |
| **Content** | **Page no** |
| INTRODUCTION | 3 |
| RATIONALE | 5 |
| OBJECTIVES | 6 |
| LITERATURE REVIEW | 7 |
| METHODOLOGY | 10 |
| FEASIBILITY STUDY | 11 |
| FACILITIES REQUIRED | 12 |
| EXPECTED OUTCOMES | 13 |
| REFERENCES | 14 |

**INTRODUCTION:**

As We all know that  **ai assistant** is a digital assistant that uses **voice recognition**, language processing algorithms, and voice synthesis to listen to specific voice commands and return relevant information or perform specific functions as requested by the user.

Based on specific commands, sometimes called intents, spoken by the user, voice assistants can return relevant information by listening for specific keywords and filtering out the ambient noise.

While voice assistants can be completely software based and able to integrate into most devices, some assistants are designed specifically for single device applications, such as the Amazon Alexa Wall Clock etc

Our project is also based on voice assistant, and it has got its own name **ZEN**. The voice assistant we made uses machine learning and deep learning technology to return relevant information or perform specific functions as requested by the user. Voice assistants use Artificial Intelligence and Voice recognition to accurately and efficiently deliver the result that the user is looking for. While it may seem simple to ask a computer to set a timer, the technology behind it is fascinating and complex.

**Voice recognition** works by taking an analog signal from a user’s voice and turning it into a digital signal. After doing this, the computer takes the digital signal and attempts to match it up to words and phrases to recognize the user’s intent. To do this, the computer requires a database of pre-existing words and syllables in any given language to be able to closely match the digital signal with. Checking the input signal with this database is known as **pattern recognition** and is the primary force behind voice recognition.

The main technology which plays a big role or the technology around which our project revolves is **Machine learning**. Machine learning refers to the subset of Artificial Intelligence where programs are created without the use of human coders manually creating the program. Instead of writing out the complete program on their own, programmers give the AI “patterns” to recognize and learn from and then gives the AI large amounts of data to sift through and study. So instead of having specific rules to abide by, the AI searches for patterns within this data and uses it to improve its already existing functions. One way machine learning can be helpful for Voice AI, is by feeding the algorithm hours of speech from various accents and dialects.

While traditional programs require an input and rules to develop an output, machine learning tools are given an input and an output and use that to create the program itself. There are two approaches to machine learning, supervised learning, and unsupervised learning. In supervised learning, the model is given data that is already partly labelled, this means some of the data given will be already tagged with the correct answer. This helps guide the model into categorizing the rest of the data and developing a correct algorithm. In unsupervised learning, none of the data is labelled, so it is up to the model to find the pattern correctly. One of the reasons this is very useful is because it allows the model to find patterns that the creators might have never found on their own, but the data is much more unpredictable.

Today many conversational assistants today combine both a task-oriented and knowledge-oriented workflow to carry out almost any task that a user can throw at it. A **task-oriented** workflow might include filling out a form, while a **knowledge-oriented**workflow includes answering what the capital of a state might be or specifying the technical specifications of a product. A **task-oriented approach**is using goals to tasks to achieve what the user needs. This approach often integrates itself with other apps to help complete tasks. For example, if you were to ask a voice assistant to set an alarm for 3PM, it would understand this to be a task request and communicate with your default Clock application to open and set an alarm for 3PM. It would then communicate with the app to see if anything else was necessary, such as a name for the alarm, then it would communicate this need back to you. This approach does not require an extensive online database, as it is mainly using the knowledge and already existing skills of other installed applications. A **knowledge-oriented approach**is the use of analytical data to help users with their tasks. This approach focuses on using online databases and already recorded knowledge to help complete tasks. An example of this approach is anytime a user asks for an internet search, it will use the online databases available to return relevant results and recommend the highest search result. If someone is searching up a trivia question, this would use a knowledge-oriented approach as it is searching for data instead of working with other apps to complete tasks. **Our project is based on task-oriented approach.**

**RATIONALE**

There are many use cases for using a voice assistant in todays’ world. For example, when your hands are full and you are unable to use a touch screen or keyboard, or when you are driving Let’s say you are driving and you need to change your music, you could just ask a voice assistant, “play my driving playlist”. This leads to a safer driving experience and helps avoid the risk of distracted driving.

They can perform generalized functions but might not perform complex and advanced tasks. These assistants are simple to use and can understand simple voice commands. **Saves Time by Automating Repetitive Tasks** Automating repeated tasks to a voice-activated personal assistant frees up the human time and resources. Also, it can efficiently perform these mundane tasks with no errors, which often leads to an improvement in customer satisfaction.

**OBJECTIVES**

### **Improved Customer Engagement**

### **Advanced Search Capabilities**

### **Smart Working Environment**

### **Let Go of The Language Barrier**

**LITERATURE REVIEW**

**Research Paper 1**

Title: Scaling Up Machine Learning: Introduction[1]

Author: Ron Beckerman

• Distributed and parallel processing of very large datasets has been employed for decades in specialized, high-budget settings, such as financial and petroleum industry applications.

• The current rise in interest in scaling up machine learning applications can be partially attributed to the evolution of hardware architectures and programming frameworks that make it easy to exploit the types of parallelism realizable in many learning algorithms.

• Several platforms make it convenient to implement concurrent processing of data instances or their features. This allows straight forward parallelization of many learning algorithms that view input as an unordered batch of examples and aggregate isolated computations over each of them.

• Increased attention to large-scale machine learning is also due to the spread of very large datasets across many modern applications. Such datasets are often accumulated on distributed storage platforms, motivating the development of learning algorithms that can be distributed appropriately.

Finally, the proliferation of sensing devices that perform real-time inference based on high-dimensional, complex feature representations drives additional demand for utilizing parallelism in learning centric applications.

**Research Paper 2**

Title: Data Structures[2]

Author: Jean-Daniel Bissonnet

* Data structures are the keystone on which all algorithmic techniques rely. The definition of basic yet high-level data structures, with precise features and a well-studied implementation, allows the designer of an algorithm to concentrate on the core issues of the problem
* Throughout this book, we describe data structures especially designed for representing geometric objects and dealing with them. But computational geometers also make extensive use of data structures that represent subsets or sequences of objects.
* These structures can be used directly by the algorithms or modified and augmented for geometric use. The first part of this chapter recalls the terminology and features of each basic data structure used in this book. It is useful to know how these structures can be implemented and what their performances are.
* The most delicate problem is undoubtedly the one addressed by dictionaries and priority queues, which treat finite subsets of a totally ordered set (the universe). To achieve better efficiency, these structures are usually encoded as balanced binary trees.
* Finally, when the universe is finite, dictionaries and priority queues can be even more efficiently implemented by other more sophisticated techniques, the characteristics of which are given without proof in the third part of this chapter.

**Research Paper 3**

Title: Introduction to Python[3]

Author: Jaan Kiusalas

* This chapter is not a comprehensive manual of Python. Its sole aim is to provide sufficient information to give you a good start if you are unfamiliar with Python
* Python is an object-oriented language that was developed in the late 1980s as a scripting language (the name is derived from the British television show Monty Python's Flying Circus). Although Python is not as well known in engineering circles as some other languages, it has a considerable following in the programming community
* In fact, Python is used by more programmers than Fortran. Python may be viewed as an emerging language because it is still being developed and refined. In the current state, it is an excellent language for developing engineering applications - Python's facilities for numerical computation are as good as those of Fortran MATLAB.
* Python programs are not compiled into machine code but are run by an interpreter. The great advantage of an interpreted language is that programs can be tested and debugged quickly, allowing the user to concentrate more on the principles behind the program and less on programming itself.
* Because there is no need to compile, link, and execute after each correction, Python programs can be developed in a much shorter time than equivalent Fortran or C programs.

**Research Paper 4**

Title: Python- A Comprehensive yet Free Programming Language for statisticians[4]

Author: Xitij U. Shukla & Dinesh J.Parmar

* Due to availability of tools for efficient numerical and statistical analysis and visualization of voluminous data with least programming syntaxes.
* Python being a modular programming language is accepted by academic and scientific community on account of its clean and readable coding style. Code indented with whitespace for function definition, loop, and control structures rather than curly braces of C, C++, or Java.
* Python is a platform independent programming language available for windows and Linux. It can run with 500MHz processor and 256 MB RAM. The manual installation may distract the researcher.
* Packages such as Pandas and Seaborn compensate further needs of advanced data analysis and visualization respectively for financial and statistical computations.

**Research Paper 5**

Title- Data Science: Fundamental Principles[5]

Author: - Alisa Bilal Zorić

* With the enormous increase in data, there is a constant need for analyzing such a large amount of data. Data Science can manage this data and develop beneficial machine learning models that predict future results.
* We can conclude that Data Science is emerging multidisciplinary field with roots in mathematics, statistics, and computer science.
* The main goal of Data Scientists is to recognize and use meaningful insights from data to help organizations in taking smarter decisions.
* During that process, they use different tools and methods to identify redundant patterns and hidden knowledge within the data. They also use the most powerful hardware, most efficient algorithms, and programming systems to solve the data related problems.

**METHODOLOGY**

**WHATSAPP AUTOMATION: -**In simple terms, WhatsApp Automation is **the use of the WhatsApp Business App or WhatsApp Business API to automate conversations with customers**. WhatsApp automated messages are pre-set replies to messages received from new or existing customers on WhatsApp.

Here we just have to mention the person’s name with the help of speech and the assistant will automatically recognized the speech and convert into text. After that it will search the name in the database and sends the message with an appropriate time (in 24hrs format).

**YOUTUBE AUTOMATION: -**The term “YouTube automation” is a bit of a misnomer. It usually means **farming out work to freelancers rather than relying on an automated process**. It is hardly a new idea and yet one that has recently become more popular.

In YouTube automation we just say, “YouTube search” and the name of the thing we are searching then it will automatically display the content related to the topic

**SPEED TEST: -**To get the most out of automation testing, start testing early in the sprint development lifecycle. Run tests as often as required. By doing so, testers can start detecting bugs as they appear and resolve them immediately.

We just command the assistant about the internet speed, and it will describe you about the uploading and the downloading speed.

**TEMPRATURE*: -*** In temperature it is used to find the temperature, speech is converted to text and then with the help of online dataset and other important parameters it is used to convert the temperature and find the result in standard form.

**MOODLE/ANOTHER APPS: *-*** Some of these apps are accessed on regular basis such that there must be alternative so that we can access them without any difficulty. Here also we can access it with the help of speech.

**TRANSLATION: *-*** an **act, process, or instance of translating**: such as. a: a rendering from one language into another also: the product of such a rendering. b: a change to a different substance, form, or appearance: conversion

Here we just need to specify an information in a language, and it will automatically display the original language to the desired language.

**Some of the important automations which are also use:-**Video downloader, Screenshot**,**Take Hindi and Open other important applications*.*

**FEASIBILITY STUDY**

COMMAND Module

The commands module **contains utility functions for working with shell command output under Unix**. This module is made obsolete by the subprocess module. There are 3 functions in the commands module for working with external commands. The functions are shell-aware and return the output or status code from the command.

PYTTSX3 Module

**pyttsx3** is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. An application invokes the pyttsx3.init() factory function to get a reference to a pyttsx3. Engine instance. it is a very easy to use tool which converts the entered text into speech. The pyttsx3 module supports two voices first is female and the second is male which is provided by “sapi5” for windows

PYWHATKIT Module

[PyWhatKit](https://pypi.org/project/pywhatkit/) is a Python library with various helpful features. It's easy-to-use and does not require you to do any additional setup. Currently, it is one of the most popular libraries for WhatsApp and YouTube automation. New updates are released frequently with new features and bug fixes.

PYAUTOGUI Module

Python pyautogui library is an automation library that allows mouse and keyboard control. Or we can say that it facilitates us to automate the movement of the mouse and keyboard to establish the interaction with the other application using the Python script. It provides many features,

* We can move the mouse and click in the other applications' window.
* We can send the keystrokes to the other applications. For example - filling out the form, typing the search query to browser, etc.
* We can also take snapshots and give an image.
* It allows us to locate a window of the application, and move, maximize, minimize, resizes, or close it.
* Display alert and message boxes.

**FACILITIES REQUIRED**

MICROSOFT VISUAL STUDIO CODE

Visual Studio Code is **a streamlined code editor with support for development operations like debugging, task running, and version control**. It aims to provide just the tools a developer needs for a quick code-build-debug cycle and leaves more complex workflows to fuller featured IDEs, such as Visual Studio IDE.

Visual Studio Code is a lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux. It comes with built-in support for JavaScript, TypeScript and Node.js and has a rich ecosystem of extensions for other languages and runtimes (such as C++, C#, Java, Python, PHP, Go, .NET).

**EXPECTED OUTCOME**

* to return relevant information or perform specific functions as requested by the user
* helps to predict accuracy of the various datasets
* translates the language
* searches for the websites and opens them

**REFERENCES**

[1] Scaling up Machine Learning Algorithms for Large Datasets

[2] Data Structures and Algorithms in Pen-based Computing Environments

[3] Introduction to Python

[4] Python-A comprehensive yet free programming language for Statisticians

[5] Fundamentals of Data Science for Future Data Scientist

[6] Got idea from [BOB Microsite Online Contest at Techgig.com](https://www.techgig.com/challenge/bank-of-baroda-hackathon-2022?sourcetype=camp617)

[7] Used Wikipedia for relevant information regarding machine learning algorithms and voice assistant

[8] Used information from blogs of Analytics Vidhya