Voice Assistant

High Level Design

Domain: Deep Learning

Creator: Vinayaka Uppar

Date: 07.12.2023

Document Version Control

Date issued	Version	Description	Author
DEC 03, 2023	1.1	First Draft	Vinayaka.Uppar
DEC 05, 20223	1.2	Added unit test cases	Vinayaka.Uppar
DEC 07, 2023	1.3	Correcting font sizes	Vinayaka.Uppar

Contents

Abstract

The abstract voice assistant is a cutting-edge technology designed to enhance user interaction with digital devices through natural language commands. In its Home Screen, users encounter a minimalistic interface featuring a voice input button for initiating commands, a display of recent commands/history, and an access point to additional settings or help options. This design prioritizes user convenience and simplicity.

Upon navigating to the Voice Input Screen, users are presented with key components such as a microphone icon, a designated area for voice input, and intuitive cancel/submit buttons. This screen serves as the entry point for users to communicate their commands vocally. The design focuses on creating a seamless and user-friendly experience during the input phase.

During command processing, the Processing Screen engages users with a loading spinner and processing message, indicating that the system is actively working on interpreting and executing the provided voice command. This visual feedback reassures users that their input is being acknowledged and processed.

Introduction

What is High-Level Design Document?

The goal of this HLD or a high-level design document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of design aspects and define them in detail
- Describe all user interfaces being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and architecture of the project

- List and describe the non-functional attributes such as security, reliability, maintainability, portability, reusability, application compatibility. resource utilization, serviceability

Scope

The HLD documentation presents the structure of the system, such as database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly technical terms which should be understandable to the administrators of the system.

General Description

Definitions

Term	Description
NLP	Voice assistant
Database	Collection of the Information
Cloud	A data center full of services connected to the internet performing service
IDE	Integrated Development Environment
UI	User Interface
API	API for interacting with server
JSON	JSON for data wrapping

Product Description

Natural Language Processing (NLP) Models:

Transformer Models: Transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and T5 (Text-to-Text Transfer Transformer), have been applied to NLP tasks. These models can understand context, semantics, and generate human-like responses.

Problem Statement

A voice 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.

It should be capable of voice interaction, music playback,

Proposed solution

Develop an advanced voice assistant using cutting-edge deep learning models for Natural Language Processing (NLP). The aim is to create a user-centric, conversational experience that understands and responds to natural language voice commands.

Further improvements

The Intelligent Voice Assistant (IVA) can seamlessly integrate into any website or application, providing users with quick responses through a friendly user interface. This embedded functionality enhances accessibility and user engagement across various digital platforms.

Data requirements

Data requirement completely depend on our problem statement. We need the dataset from:

- Date time. Weekday, month can be extracted from it.
- Weather condition like cloudy, rain, smoke and etc.
- Temperature of weather on specific date time.
- Voice commands ,contextual information
- User patterns

These are the required parameters to feed into model.

Tools used

Python programming language and frameworks such as NLP,TTS,(IDE) vscode,github and a few other libraries were used to build the whole model.

- Google Cloud Speech-to-Text: Accurate speech-to-text conversion.
- Dialogflow: Natural Language Processing (NLP) for user intent understanding.
- Flask: Lightweight web framework for application and server development.
- Google Cloud Text-to-Speech: Converts text responses into natural-sounding speech.
- GitHub: Version control system for collaborative development and code management.

- Apache Cassandra: Distributed and scalable database for data storage and retrieval.

Hardware Requirements

- Windows 7 8 11
- Minimum 1.10 GHz processor or equivalent.
- Between 1-2 GB of free storage
- Minimum 512 MB of RAM
- 8 GB of hard-disk space

Constraints

The front-end must be user friendly and should not need any one to have any prior knowledge in order to use it.

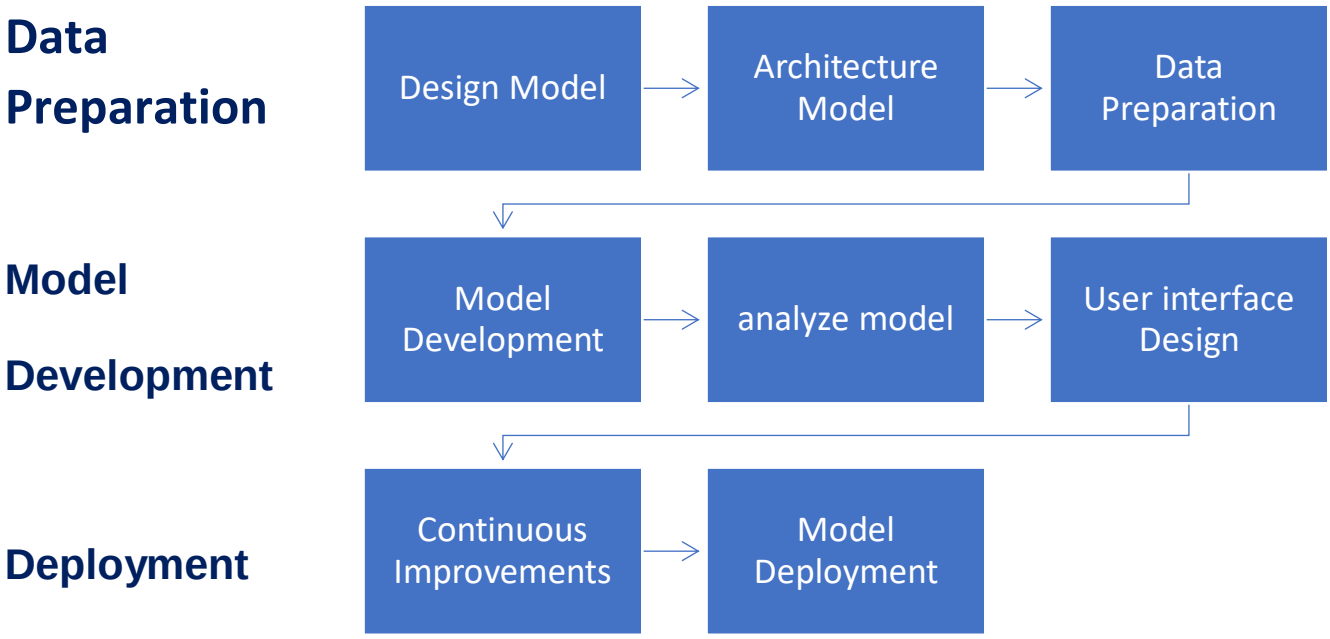
Assumptions

The main objective of this project is to implement the use case as previously mentioned (2.3 problem statement) for new dataset that comes through the UI. It is assumed that all aspects of this project have the ability to work together as the designer is expecting and also the data on which our model is trained is as correct as possible

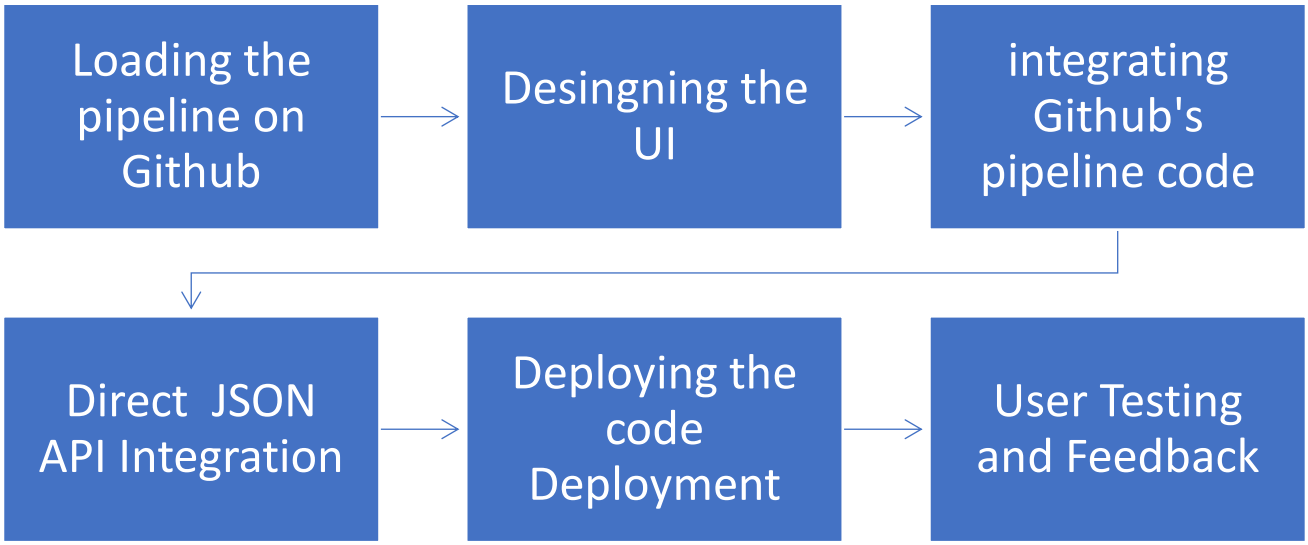
Design Details

Process Flow

For accomplishment of the task, we will use a trained Deep Learning model. The process flow diagram is shown below:



Deployment



Event Log

The system should log every event so that the user will know what process is running internally. Initial step-by-step description: 1. The system identifies at what level logging is required 2. The system should be able to log each and every system flow 3. Developer can choose logging method. You can choose database logging/ File logging as well 4.

System should not hang even after so many loggings. Logging just because we can easily debug issues, so logging is mandatory to do.

Error Handling

Errors should be encountered, an explanation will be displayed as to what went wrong ? An error will be defined as anything that falls outside the normal intended usage.

Performance

Natural Language Processing (NLP) tool used plays a crucial role in understanding and interpreting user commands. Here are some popular NLP tools that could be utilized for such a project

Reusability

The code written and the components used should have the ability to be reused with no problems.

Application Compatibility

The different components for this project will be using Python as an interface between them, each component will have its own task to perform, and it is the job of Python to ensure proper transfer of information.

Resource Utilization

When any task is performed, it will likely use all the processing power available to it until finished.

Dashboards

While not a traditional NLP tool, Wolfram Alpha is a computational knowledge engine that can be integrated for understanding and responding to factual queries in a voice assistant.

KPIs (Key Performance Indicators)

- Key Performance Indicators of NLP
- Latency or the amount of Commands recognize by voice assistant
- The processing and analyzing various task

Conclusion

the Voice Assistant project represents a comprehensive and meticulously designed system aimed at delivering an intelligent, user-centric, and responsive conversational experience. By integrating cutting-edge technologies and tools, the project ensures robust speech recognition, accurate natural language understanding, and effective user interaction. The following key points summarize the project's significance and potential impact