

Voice-Controlled Coffee Machine with Raspberry Pi and Whisper

Lena Kittel, Alicia Diem

1 Introduction

In this project, our goal was to build a voice-controlled system to operate a coffee machine using a Raspberry Pi and the Whisper model from OpenAI. The idea was to capture spoken commands, transcribe them using the speech-to-text model Whisper, and trigger an action such as pressing a button through a servo motor. Specifically, we wanted the system to respond to the command “make coffee”.

2 Results

2.1 Milestone 1

Our first step was to select suitable hardware for audio input. We initially attempted to use an audio HAT (WM8960), but encountered compatibility issues between its driver software and the Raspberry Pi 5 OS. After considering downgrading the operating system, we opted instead for a more reliable solution: a USB microphone.

Using a Python script, we recorded audio input from the USB microphone in ‘.wav’ format. This recording was then processed by OpenAI’s Whisper model via its API. We configured the system to listen for predefined keywords or “buzz words”, and trigger specific actions accordingly.

2.2 Milestone 2

Once the audio processing pipeline was working, we planned to integrate a servo motor to simulate a button press on the coffee machine. Unfortunately, we were unable to complete this part of the project. Although we researched and tested several Python libraries for controlling servo motors, the scripts we tried either did not function properly or were not compatible with our setup. As a result, the action following the recognized “make coffee” command could not be executed through hardware control.

2.3 Highlight

The highlight of the project was implementing the full voice recognition and transcription pipeline using Whisper. Being able to record and transcribe voice commands marked an important milestone, even though the hardware control could not be finalized.

3 Conclusions

Despite several technical challenges, especially with audio hardware compatibility and servo motor control, we were able to implement a working voice interface on the Raspberry Pi using OpenAI's Whisper API. While the final hardware actuation step could not be completed due to script issues, the project offered valuable insights into speech recognition, Python scripting, and system integration.