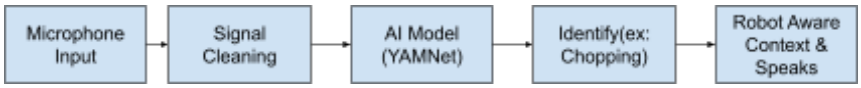
Team members: Joshua Kirby, Alan Nur
Instructor: Luyao Niu
TECHIN 513 A, 2026-01-11

Project Proposal

Problem statement

Oftentimes, people leave the stove on or sink running, a toilet takes too long to refill or the AC makes strange noises. There are so many audio signals within a given living space. We plan to improve in-home, AI voice assistants by bringing them a clearer understanding of a household and its residents through the expansion of its data models into the realm of menial, domestic sounds,[3] and adding the capability for assistants to remedy potential home living inconveniences.



Proposed methodology

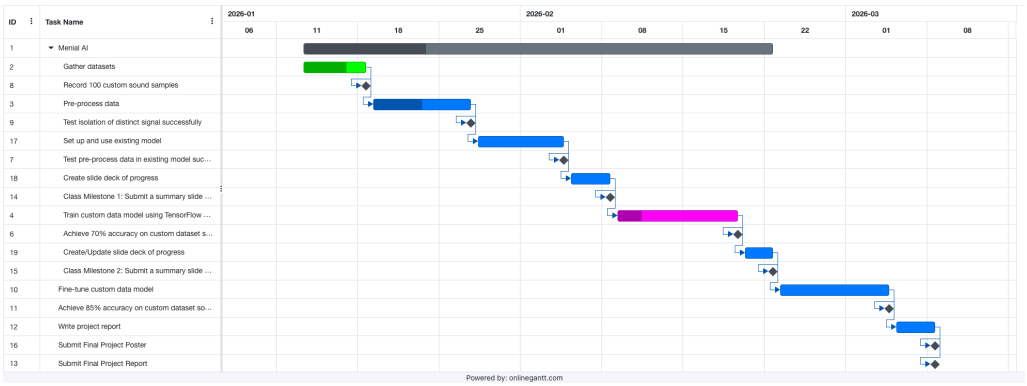
Our current plan is to use four techniques to gather all relevant information from dataset samples. We'll use (1) MFCC for capturing the sample's timbral and spectral characteristics, (2) STFT for analyzing spectrographic overlapping frequency changes overtime, (3) ZCR for separating and categorizing sounds by their onset times, and (4) NMF for approximating separate sound sources (see Acronyms section below). The datasets we plan to use include AudioSet and ESC-50. In addition, we will build a custom dataset of a few dozen distinct, domestic sounds, recorded on different mics in both isolated and natural scenarios.

Expected outcomes and deliverables

Our goal is to build a working prototype on a laptop or Raspberry Pi that listens for chopping, boiling water, and running taps. When it hears one of these sounds, it will immediately ask a helpful question, like "Do you need the next recipe step?" We will consider the project a success if the system identifies the sounds correctly 85% of the time, responds in less than two seconds, and works reliably even when there is background noise in the room.

Challenges and milestones

A major challenge is that some sounds, like frying food and heavy rain, are very similar. We will solve this by training the model specifically on these confusing pairs so it learns the small differences. Another issue is that kitchens are noisy with fans and talking, so we will add recorded background noise to our training data to teach the robot to ignore it. Finally, running AI on a small Raspberry Pi can be slow, so we will use a compressed "quantized" version of the model to make sure it responds instantly.



Acronyms

NMF … Non-Negative Matrix Factorization
MFCC … Mel-Frequency Cepstral Coefficient
STFT … Short-Time Fourier Transform
ZCR … Zero-Crossing Rate

References

1. Claude Sonnet 4.5. “Signal Processing Methods for Household Sound Classification.” *Claude.ai*, claude.ai/chat/5e1b23d7-7d92-4f77-8ecb-a4166add8391. Accessed 12 Jan. 2026.

    Prompt:

    I need you to research deeply into a subject for me. I have a final project assignment concerning managing data and signal processing. I need to be able to use Python to process some signal in a way that's no novel or not implemented in a certain way.

    You do not need to help me with the novel approach, as you are an AI model, and as I understand, it incapable of truly novel approaches. I instead need you to help me figure out what method(s) of signal processing I should use.

    For our final project, we plan on using a single microphone to gather information about ambient or manual noises coming from a residential household. This can be any sound that you'll find in someone's house. We will then take that audio signal, separate distinct noises or sound signatures, and trained an AI model to understand what each of those audio signals are.

    We mainly plan to use household voice assistance to gather data or take action based on the sounds that they're hearing.

    You must focus on researching this topic and determining at least five forms of signal processing, algorithms, or methodologies that me and my teammate could study our ourselves and implement into our Python code. DO NOT write any actual Python code implementation for me. Obviously, if you want to give example examples in Python, you can, but you don't need to do that either. The main thing for you to do is to deeply research this topic. Only present information to me if you have clear citations and resources and you are at minimum 95% certain that the information you are presenting me is true and scientifically accurate.

    I want you to make no assumptions about what the correct methodologies are. Instead commence in-depth research on all possible methodologies that I could use for this project. After your research, present me with at least five that you believe are relevant to my final project.

    DO NOT make assumptions about what the correct methodologies are. Research what relevant mythologies are out there.

    Begin your in-depth research now.

2. Gemini 3 Pro. “Gemini Chat.” *Gemini.google.com*, gemini.google.com/app/a65a86ff0da1cb57. Accessed 12 Jan. 2026.

    Prompt:

    I need you to research deeply into a subject for me. I have a final project assignment concerning managing data and signal processing. I need to be able to use Python to process some signal in a way that's no novel or not implemented in a certain way.

You do not need to help me with the novel approach, as you are an AI model, and as I understand, it incapable of truly novel approaches. I instead need you to help me figure out what method(s) of signal processing I should use.

For our final project, we plan on using a single microphone to gather information about ambient or manual noises coming from a residential household. This can be any sound that you'll find in someone's house. We will then take that audio signal, separate distinct noises or sound signatures, and trained an AI model to understand what each of those audio signals are.

We mainly plan to use household voice assistance to gather data or take action based on the sounds that they're hearing.

You must focus on researching this topic and determining at least five forms of signal processing, algorithms, or methodologies that me and my teammate could study our ourselves and implement into our Python code. DO NOT write any actual Python code implementation for me. Obviously, if you want to give example examples in Python, you can, but you don't need to do that either. The main thing for you to do is to deeply research this topic. Only present information to me if you have clear citations and resources and you are at minimum 95% certain that the information you are presenting me is true and scientifically accurate.

Begin your in-depth research now.

3. Various authors. “Learn about Sound Detection for Speakers & Displays - Google Nest Help.” *Google.com*, 2019, support.google.com/googlenest/answer/9681286.

4. Various authors. “Sound Detection Features Not Available?” *Googlenestcommunity.com*, 29 May 2025, www.googlenestcommunity.com/t5/Subscriptions-and-Services/Sound-Detection-Features-Not-Available/m-p/720144. Accessed 12 Jan. 2026.

5. Bosch Global. “Audio AI: Learning to Understand Sounds.” *Bosch Global*, 2024, www.bosch.com/research/research-fields/artificial-intelligence/audio-ai/.

6. Lindahl, Mallory. “Acoustic Awareness: ANAVI Framework Addresses Robot Noise Levels in the Home.” *Cmu.edu*, 28 Jan. 2025, www.cmu.edu/news/stories/archives/2025/january/acoustic-awareness-anavi-framework-addresses-robot-noise-levels-in-the-home. Accessed 12 Jan. 2026.

7. Kosmin, Danylo. ““MeowTalk” — How to Train YAMNet Audio Classification Model for Mobile Devices.” *AKVELON*, 21 Sept. 2020, akvelon.com/meowtalk-how-to-train-yamnet-audio-classification-model-for-mobile-devices/. Accessed 12 Jan. 2026.

8. TensorFlow. “Transfer Learning with YAMNet for Environmental Sound Classification | TensorFlow Core.” *TensorFlow*, 16 Aug. 2024, www.tensorflow.org/tutorials/audio/transfer_learning_audio.

9. Chollet, François. “Transfer Learning & Fine-Tuning.” *Keras.io*, 15 Apr. 2020, keras.io/guides/transfer_learning/.

10. Gemmeke, J., et al. "Audio Set: An ontology and human-labeled dataset for audio events." *Google Research*, 2017. https://research.google.com/audioset/ Google Research. "Domestic sounds - AudioSet Ontology." Accessed Jan 2026.

11. Sound and Video Understanding teams. “Domestic Sounds, Home Sounds.” *AudioSet*, 2026, research.google.com/audioset/ontology/domestic_sounds_home_sounds_1.html. Accessed 12 Jan. 2026.

12. Piczak, Karol, and Various authors. “ESC: Dataset for Environmental Sound Classification.” *GitHub*, 6 Apr. 2018, github.com/karolpiczak/ESC-50.

13. Artem Kirsanov. “Wavelets: A Mathematical Microscope.” YouTube, 15 Aug. 2022, www.youtube.com/watch?v=jnxqHcObNK4.