

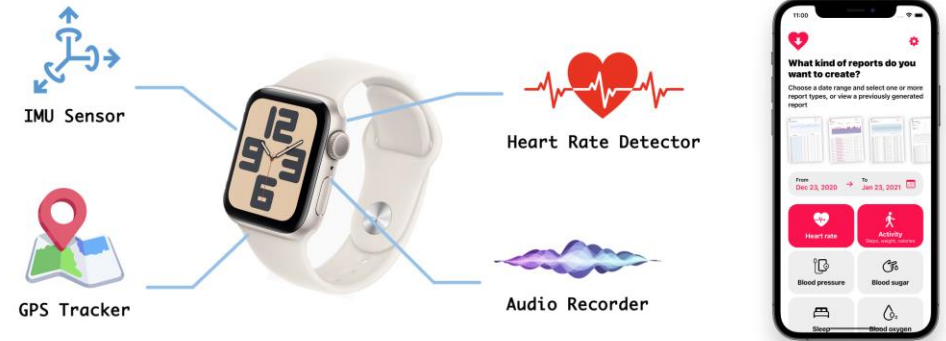
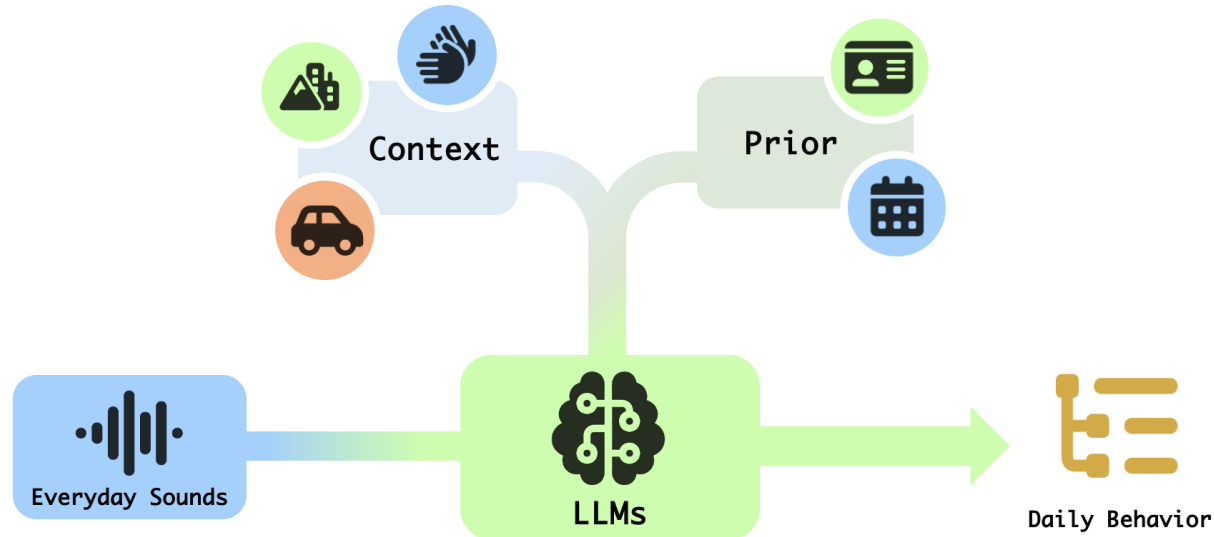
Behavior Prediction from Everyday Sounds via LLMs with Multi-sensor Context and Priors

Intro:

Sounds are ubiquitous in daily life: wind while running, clinking at meals, keyboard taps at work. With some context and prior knowledge, people can imagine what's happening just by sounds.

Yet machines can't — until now. By leveraging everyday devices like smartwatches, we can capture real-time audio and sensor data. Sensor readings help infer context such as location and activity, which, combined with user habits, form auxiliary knowledge. This knowledge, along with raw audio, is fed into a multimodal large model to predict the user's current behavior and state.

With continuous recordings, we can analyze users' daily routines and provide detailed daily activity lists and behavior reports for self-reflection and understanding.



Dataset:

Given the specificity of data required, no public dataset fits our project needs. Therefore, we will construct a custom dataset via smartwatches, mainly consisting of:

- User Questionnaire: Gathers basic demographics and routine behaviors as prior knowledge.
- IMU Data: Captures motion signals to identify specific limb movements.
- GPS Location: Provides spatial data to infer environment and mobility states (e.g., walking, running, commuting).
- Heart Rate: Records physiological responses to assist in activity classification.
- Raw Audio: Continuously records sound for behavioral inference by the model.

Author: Xingjian Tian, Fangfei Gou, Xin Tang, Zeyu Wang, Yingke Ding