An AI-Based System Utilizing IoT-Enabled Ambient Sensors and LLMs for Complex Activity Tracking

Yuan Sun, Rutgers University
Jorge Ortiz, Rutgers University

Presenter: Ko Wei Su
National Cheng Kung University

國立成功大學
National Cheng Kung University

Outline

- **Introduction**
- **Proposal**
- **Related Work**

Introduction

The paper introduces a **non-intrusive smart sensing system** leveraging large language models (LLMs) to assist in elderly care.

The system detects complex activities composed of more than two atomic activities.

Atomic activities refer to short-term, unit-level tasks captured by sensors that cannot be further divided.

Introduction

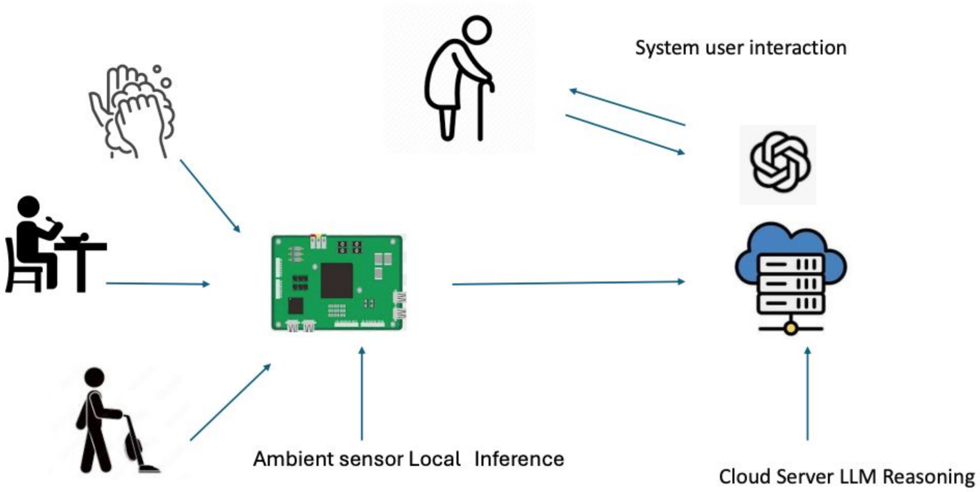


Figure 1: The system employs both local inference using ambient sensors and reasoning via a cloud-based LLM. The sensors detect atomic activities, and the cloud server receives these activity sequences as context to further detect higher-level meanings, make decisions, and interact with the user.

Sensor Board

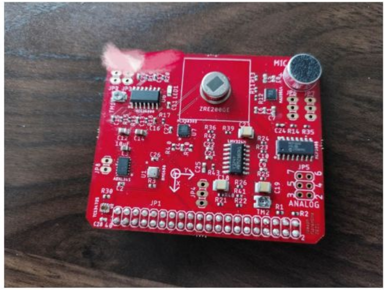


Figure 2: The non-intrusive sensor board we design for our system

Figure 3: Raspberry Pi Model B+ used in our ambient sensor setup, facilitating seamless integration for elderly care assistance and activity tracking

- Device enhances sensor data explanation while ensuring privacy using non-invasive sensors.
- Sensors include PIR (motion), IMU (accelerometers), RGB, pressure, humidity, magnetometer, gas, and temperature.

Data Collection

Figure 4: Eating activity collected by the ambient sensor

Experiments

- The model uses a channel-wise MLP for feature extraction and a sensor fusion module, also MLP-based.
- A fast Fourier convolution module was added to enhance accuracy.

Activity Name	F1	Precision	Recall
eat	1.00	1.00	1.00
paperdis	0.43	0.50	0.38
write	1.00	1.00	1.00
chop	1.00	1.00	1.00
hand wash	1.00	1.00	1.00
pour water	0.48	0.50	0.47
clean floor	1.00	1.00	1.00
knock	1.00	1.00	1.00
run	1.00	1.00	1.00
curtain	1.00	1.00	1.00
light Switch	1.00	1.00	1.00
type	1.00	1.00	1.00
door pass	1.00	1.00	1.00
wipe desk	1.00	1.00	1.00
chat	0.32	0.33	0.31
basketball	1.00	1.00	1.00
saw	1.00	1.00	1.00
shave	1.00	1.00	1.00
wash dish	1.00	1.00	1.00
teeth	1.00	1.00	1.00

Table 1: Initial results of F1 scores for each activity

Thoughts

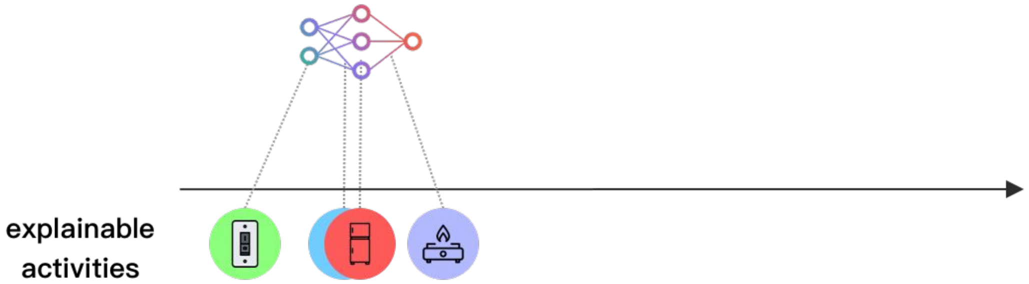
- **Proposed Framework**
- **Related Works**

Proposed Abstract

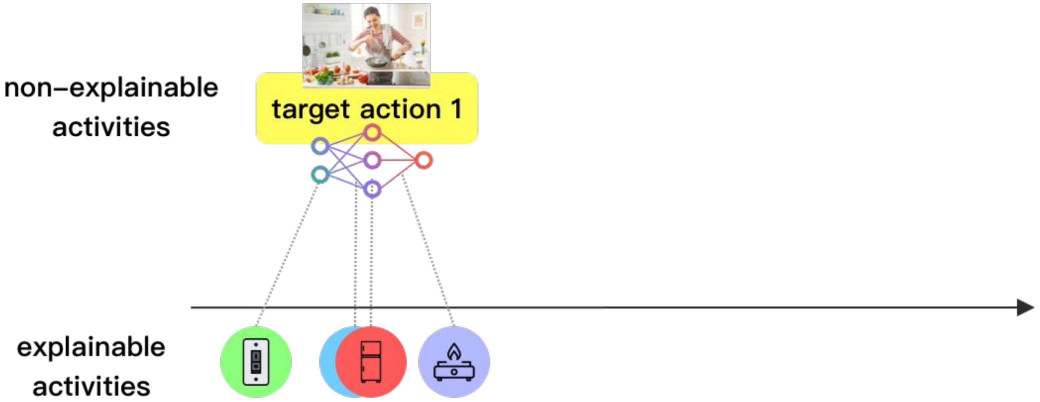
With the Healthcare IoT market projected to reach US$83.81bn globally by 2024 and growing at a CAGR of 9.91% through 2029, the integration of smart devices is becoming crucial in healthcare settings. This paper presents a sensor-based system designed to predict non-explainable actions from explainable actions, using sensor data to provide timely warnings for critical activities. By leveraging non-intrusive sensors and advanced data processing, the system anticipates potential risks, such as reminding users to turn off the stove after cooking and before leaving the house. Initial results show high accuracy for explainable actions, enhancing safety and proactive care in healthcare and elderly care environments.

Thoughts

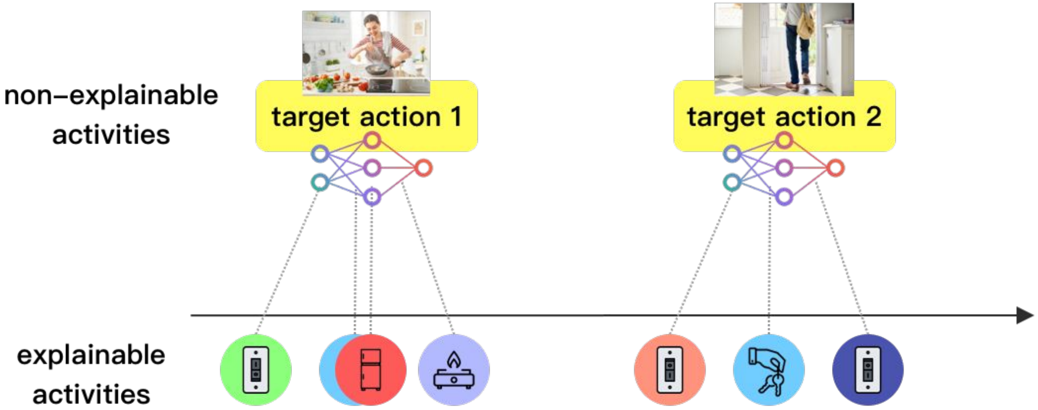
Thoughts



Thoughts

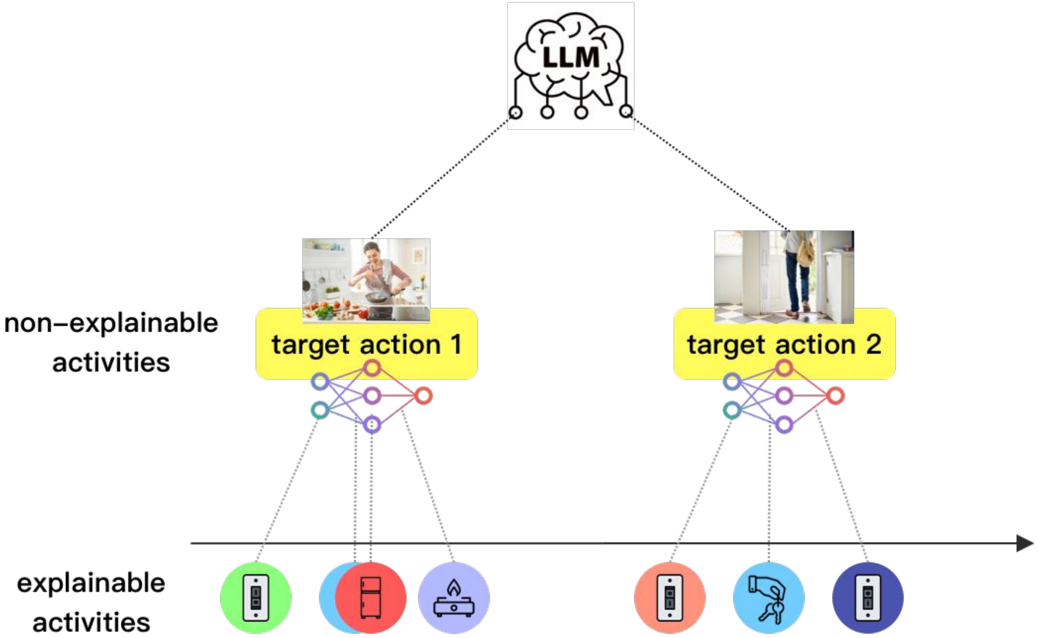

non–explainable
activities
target action 1
explainable
activities

Thoughts



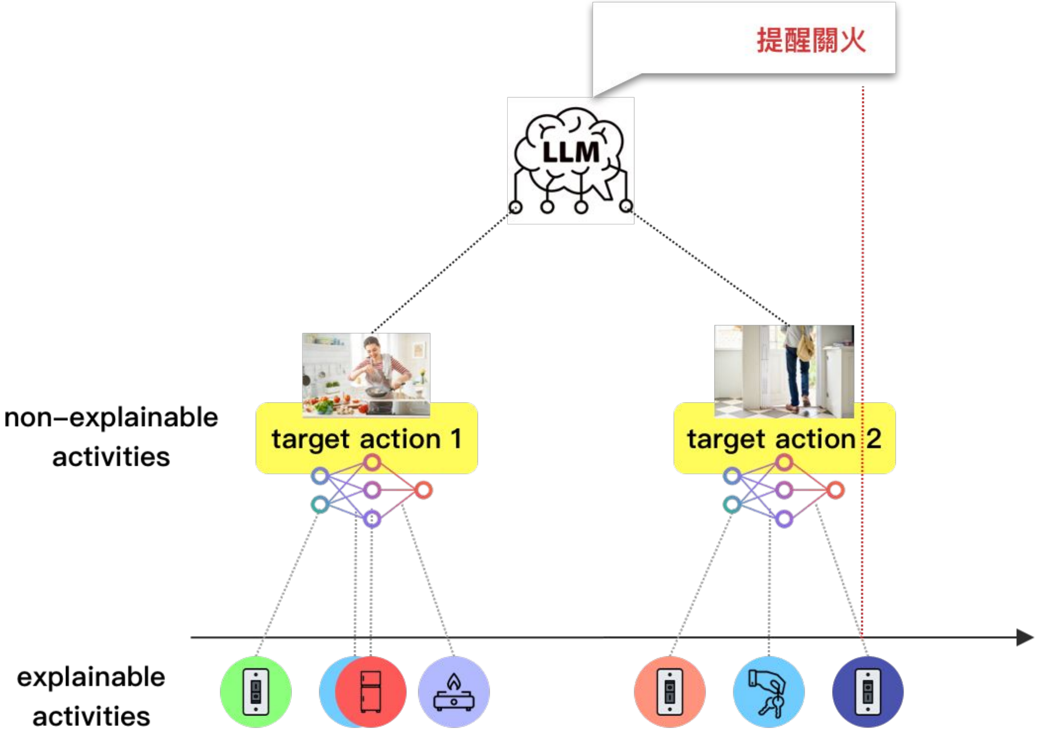
non–explainable
activities
target action 1
target action 2
explainable
activities

Thoughts



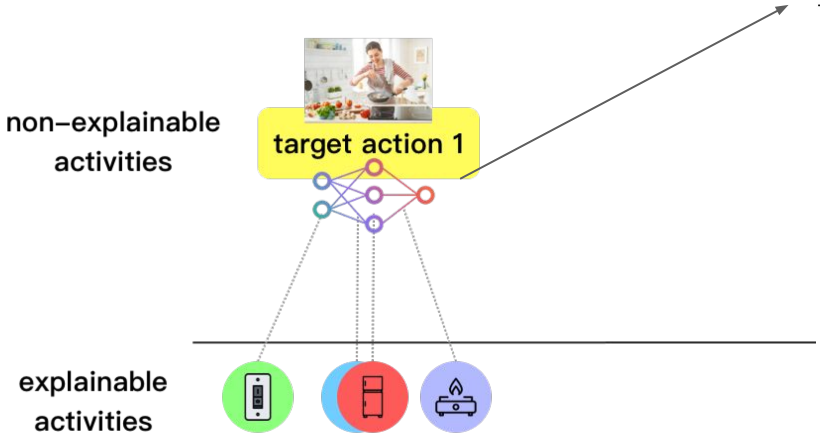
LLM
non–explainable
activities
target action 1
target action 2
explainable
activities

Thoughts


提醒關火
LLM
non–explainable
activities
target action 1
target action 2
explainable
activities

Related Works

https://dl.acm.org/doi/10.1145/3580804



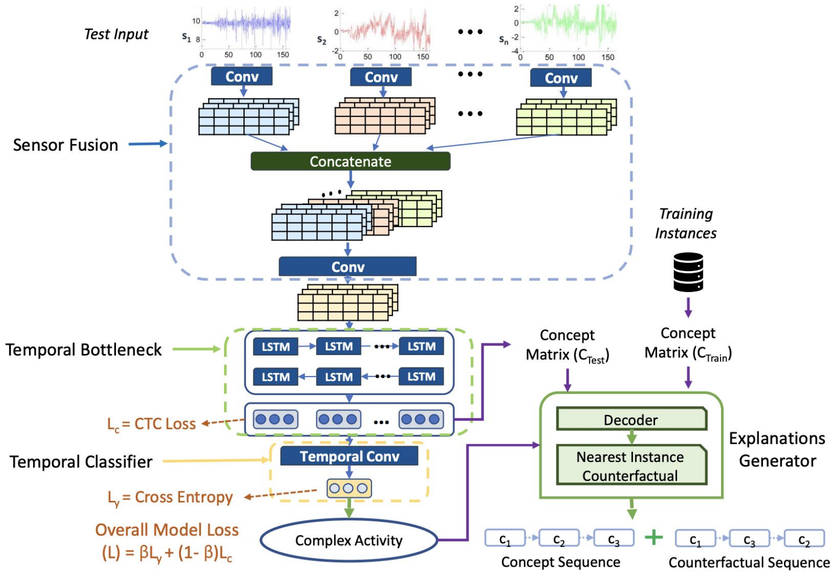


Fig. 3. The overall *X-CHAR* model design showing the various layers in each module.

Restaurant Activity	Concept Sequence	Concepts
Using Restroom (Hygienic)	Footsteps → Using toilet → Toilet flush → Wash Hands	footsteps, using toilet, flush toilet, wash hands, open shelf, chopping vegetables, peeling, using blender, take glass, pour water, drinking
	Wash Hands → Using toilet → Toilet flush → Wash Hands	
Using Restroom (Unhygienic)	Footsteps → Using toilet → Toilet flush	
	Footsteps → Wash Hands → Using toilet → Toilet flush	
	Wash Hands → Using toilet → Toilet flush → Footsteps	
Making a fruit juice	Opening shelf → Chopping → Using blender	
	Peeling → Chopping → Using blender	
	Opening shelf → Peeling → Using blender	
Making a puree/sauce	Peeling → Using blender → Chopping → Using blender	
	Chopping → Using blender → Peeling → Using blender	
	Open Shelf → Using blender → Chopping → Using blender	
Having a drink	Take glass → Pour water → Drink	

Jeya Vikranth Jeyakumar, Ankur Sarker, Luis Antonio Garcia, and Mani Srivastava. 2023. X-CHAR: A Concept-based Explainable Complex Human Activity Recognition Model.

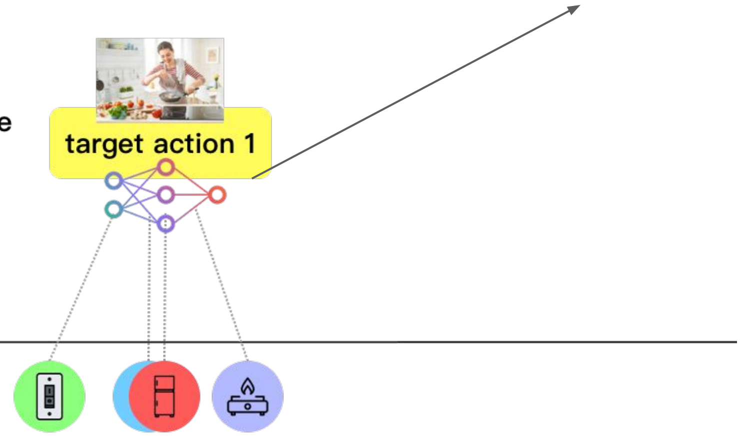
Related Works

HARGPT: Are LLMs Zero-Shot Human Activity Recognizers?

Sijie Ji*, Xinzhe Zheng* , Chenshu Wu

non–explainable activities

explainable activities



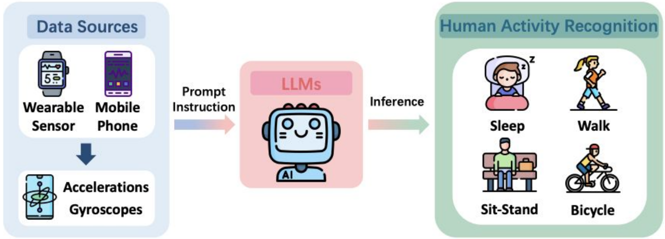


Figure 1: Workflow of HARGPT.

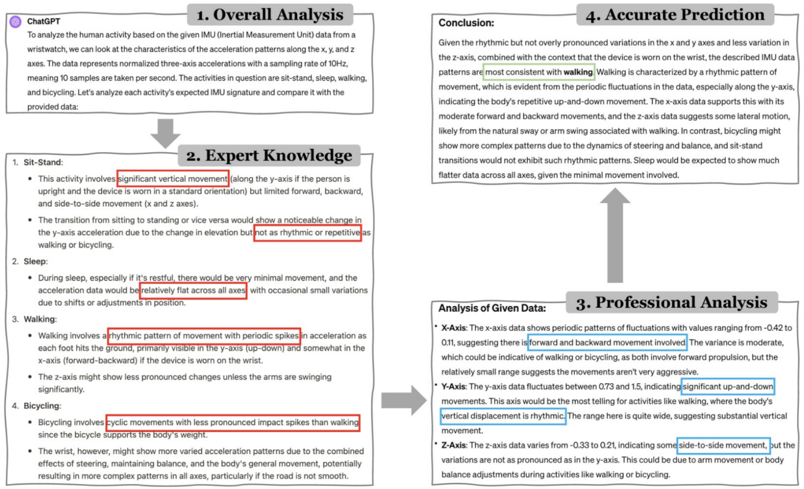


Figure 4. Detailed step-by-step inference generated by GPT4 with a walking example.