LLM-Based Activity Recognition Using AR/VR Headsets

RUTGERS
WINLAB | Wireless Information Network Laboratory

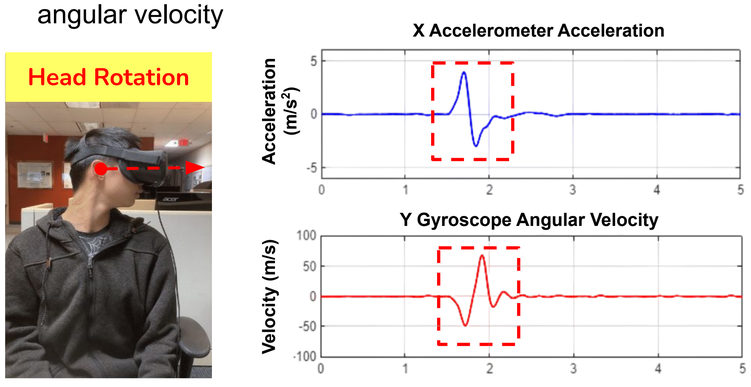
Team: Leon Du & Jerry Huang **Project Advisor:** Prof. Yingying (Jennifer) Chen
Mentors: Changming Li, Honglu Li, Tianfang Zhang

Motivation and Objective

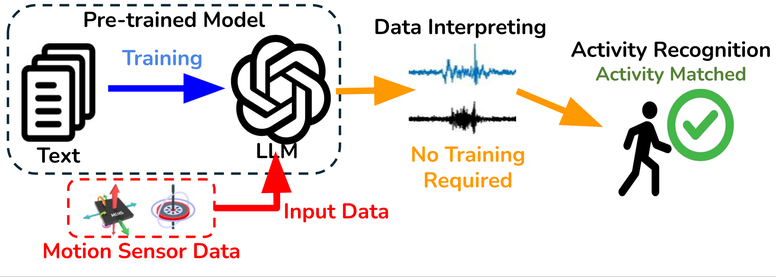
- Augmented Reality/Virtual Reality (AR/VR) technologies have been rapidly gaining popularity over the last decade [1]
- AR/VR has been utilize in various types of applications (i.e., gaming, banking, eduction, shopping, etc.s)
- User movements can be captured via built-in motion sensors (i.e., accelerometer and gyroscope) embedded in AR/VR Head Mounted Device (HMD)
- LLMs shows high generality to interpret semantic interactions from input data [2]
- LLMs could be used to interpret body motions instead of current methods of neural networks, eliminating the need for large scale data collection and training

Background

- Motion sensors (accelerometer and gyroscope) recordings capture human motions in terms of linear acceleration and angular velocity



- LLMs trained on large size of text data, can understand general context pattern, exhibit strong generalization capabilities for classification tasks
- Utilize the generalization capability of LLM to interpret motion sensor readings
- No training is required for classification tasks



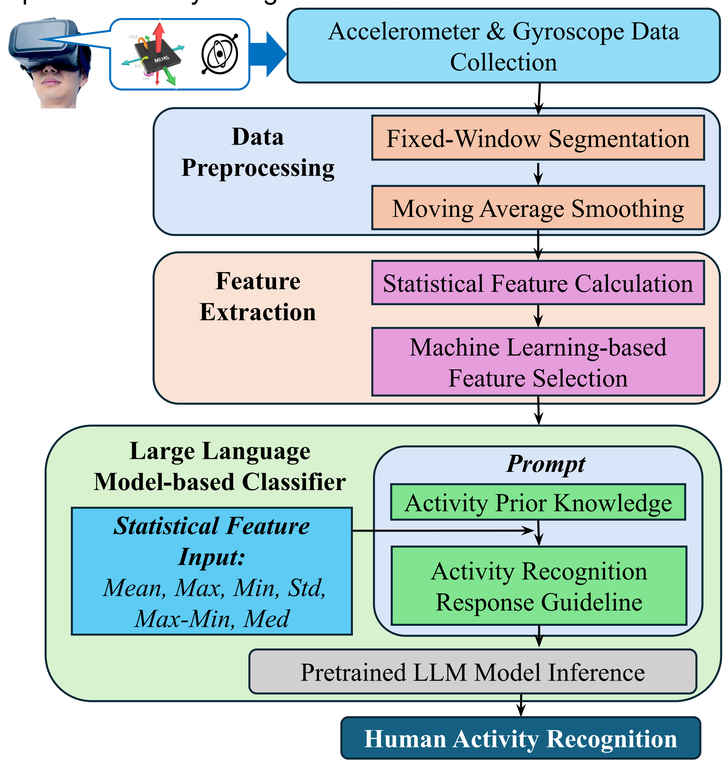
LLM Prompt Engineering

1. Explaining the goal of the task and data types to LLM
2. Extract and Identify features from the data and provide expert knowledge for how to analyze the data
3. Provide a response template for results

Sensor Features & Interpretation Guide
You will be given three types of sensor-derived features for each motion:
Y-Axis Maximum Acceleration (from Accelerometer)
- High values → Active vertical movements (e.g., stand up, walk, nod)
- Low values → Minimal vertical movement (e.g., turnleft, turnright)

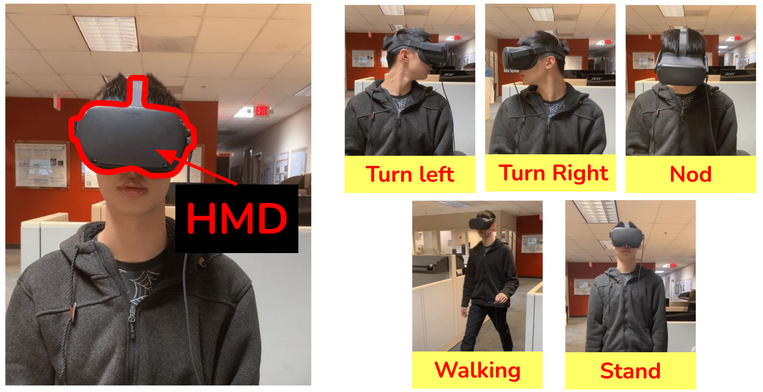
System Overview

1. **Data Collection**: Collect Accelerometer and gyroscope data during activities when wearing VR HMD
2. **Data Preprocessing**: Segment data and remove the ambient noise and outliers
3. **Feature Extraction:** Extract statistical features and use ML to select effective statistical features for classification
4. **LLM based Classifier:** Design prompts for LLM to interpret the motion data for specific activities and guide LLM to perform activity recognition task



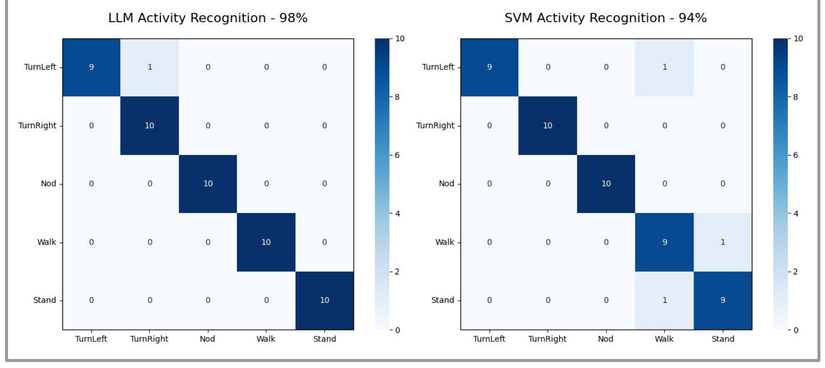
Experiment Setup

- Utilizing Android Studio, we coded an algorithm to connect the API of the motion sensors to extract motion sensor readings from the Oculus Quest's HMD
- Designed 5 motions for data collection
- Preprocessed and denoised data in MATLAB to create accurate waveforms and extract statistical features



Evaluation Results

- LLM achieves 98% recognition accuracy, outperforming the prediction accuracy of SVM



Conclusion

- With designed prompts, LLM achieves a higher accuracy compared to the baseline machine learning model SVM, indicating the potential for activity recognition without training effort from users
- With further prompt fine-tuning, activity recognition could be expanded to more complex body movements

References

[1] Zhao, Jingbo, Mingjun Shao, Yaojun Wang, and Ruolin Xu. "Real-time recognition of in-place body actions and head gestures using only a head-mounted display." In 2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR), pp. 105-114. IEEE, 2023.
[2] Guo, Xiaonan, Jian Liu, and Yingying Chen. "FitCoach: Virtual fitness coach empowered by wearable mobile devices." In IEEE INFOCOM 2017-IEEE Conference on Computer Communications, pp. 1-9. IEEE, 2017.

Acknowledgement

We would like to thank our project advisor and mentors for their support and guidance throughout this project.

WINLAB