RIDE Spring 25 Mobile LLM Development

Yi Ding, Assistant Professor

What is RIDE

If you are looking to explore collaborative design and research while earning course credit, you should consider joining one of the RIDE projects. These projects are intended to be **collaborative**, **multi-year**, **multi-generational**, **multi-disciplinary** projects that are coached by ECS faculty and instructors. These are a unique chance for you to practice being an engineer or CS professional.

There are different options depending on your level:
- ECS 1192 (2-3 hours of commitment per week)
- ECS 2192 (2-3 hours of commitment per week)
- ECS 3292 (4-6 hours of commitment per week)
- ECS 4392 (6-9 hours of commitment per week) – NOTE this requires prior enrollment in ECS RIDE courses

In this class (RIDE Spring 25 Mobile LLM Development), we will enroll a group of around **three** students and let them **work together** on a **research-oriented** project.

The Faculty and PhD Mentor

Yi Ding

Assistant Professor at CS

Research Interest:

- Mobile Computing, Spatial-Temporal Data Mining, Cyber-Physical Systems

Teaching:

- Computer Networks (Spring 24, Fall 24)
- Special Topics in CS: Data Science for Smart Cities (Fall 23)

PhD Mentor (providing guidance together with the faculty):

- Xiao Yan (xiao.yan@utdallas.edu)

Background

- Large Language Models (LLM) like ChatGPT and Perplexity are widely used given their powerful capability in world knowledge, language understanding,and reasoning.
- LLMs can play more important roles if they can run locally on mobile/edge devices with no or limited access to internet. Examples:
  - First responding in the wild.
  - Privacy-preserving APPs that does not upload data to servers.
- However, existing LLMs are too large to run efficiently on mobile devices like smartphones.

The Project

In the project, we aim to develop an APP on the smartphone that:

(1) meets specific need (e.g., healthcare, wellness, social, entertainment);
(2) collects data from different sensors (e.g., camera, microphone, GPS, IMU)
(3) utilize LLM (e.g., ChatGPT, Llama, Claude) with no or limited access to Internet;
(4) with efficient system design (i.e., latency, CPU load, memory usage, bandwidth usage)

What You will Learn and What We will Provide

In the project, you are expected to do the following:

(1) Explore different mobile applications and find an interesting APP to work on;
(2) Learn fundamental knowledge of LLM and mobile APP development;
(3) Write programs for the APP (e.g., data collection, LLM interaction, UI, system optimization) under guidance from the faculty and the PhD mentor.
(4) Write technical report and research papers;

The faculty and the PhD mentor will provide:

(1) Accessible materials on LLM and mobile development (e.g., online blogs & courses);
(2) Demo code and technical reports
(3) Weekly discussion to answer any questions you have;
(4) Shared lab space and equipment, including a 2xRTX4090 workstation and an Android phone

How Will You be Evaluated/Grade?

(1)   The final APP performance and technical report;
(2)   Weekly or monthly progress report;
(3)   Feedback from the faculty and the PhD mentor.

FAQ

Q: Any Prerequisite for the course? Do you need programming background?
A: No. If you are a CS/CE students, you are expected to be familiar with Python or other development language. If you have other background (ME/Bio), it will be better that you can provide some unique expertise.

Q: How do we collaborate?
A: The collaboration and task allocation is based on expertise and experience. The group members should talk to each other and coordinate. The faculty member and the PhD mentor will also provide feedback. You will also use version control tools like GitHub to collaborate.