CS 490 Capstone Project Proposal

Team 2: Aiman Madan, Cristian Enriquez, Jacob Almon, Kieran Hawkins.

We propose designing and developing an app called CookLit, an Android application designed to help users create recipes from available ingredients using **Qualcomm's fine-tuned YOLOv8 with food recognition** model for **object detection**. Users take photos of their ingredients, and the app identifies them via object detection. It then **retrieves tailored recipes** using a large language model (Llama), **retrieves recipe images** (API), and **retrieves historical context** (e.g., the origin of the dish). **This reduces food waste, inspires culinary creativity, and educates users about diverse cuisines**. We plan to implement the CookLit application by using a combination of **AI models and Android development tools.**

We will utilize Qualcomm's YOLO v8 model for on-device image classification, mapping identified ingredients to a structured database. Recipes will be generated using an LLM based on these ingredients. Additionally, we will use TensorFlow Lite to enable efficient, private, and real-time AI execution on mobile devices, supporting rapid cross-platform development.

The application will utilize Llama 3.2, a large language model, to generate descriptions and instructions for the recipes that were found in the database. The model will be prompted to act as an assistant for an AI ingredient app and to generate descriptions, background, and instructions for recipes we provide it. A fine-tuned prompt structure will ensure that the model generates accurate and relevant recipes, considering common pairings, dietary restrictions, and

user preferences. The generated recipes will be structured with step-by-step instructions, ingredient lists, and estimated cooking times. Users will be able to ask for elaboration on the recipe if needed.

The application will integrate the Spoonacular API to retrieve recipes based on a list of ingredients identified using Qualcomm's fine-tuned YOLOv8 model on food recognition. All the information gathered, including recipe names and a photo of the recipe will be displayed seamlessly on the user's phone.

The frontend will be developed using React Native instead of Kotlin, leveraging our team's prior experience with the framework to streamline development and ensure cross-platform compatibility across iOS and Android. While Jetpack Compose was initially considered for its modern UI capabilities, React Native's ecosystem and shared codebase advantages better align with our project's goals.

Our team is consists of:

- Aiman Madan: Driven by a passion for advancing intelligent systems, I specialize in building end-to-end solutions that bridge cutting-edge AI/ML innovation with practical applications. My expertise spans designing robust web scraping architecture for high-quality data pipelines and crafting intuitive, responsive front-end interfaces for desktop and mobile platforms.
- Cristian Enriquez: I have experience in building both mobile and web applications. In these applications I have practiced working with UIs for frontend, with databases such as MySQL for backend and with utilizing API's. I have some exposure to AI models such as TensorFlow and PyTorch that can help with this project.

- Jacob Almon: I have a strong background in web application development and backend engineering. I'm proficient in machine learning toolkits like Scikit-Learn and deep learning frameworks such as TensorFlow and PyTorch, with hands-on experience building personal projects utilizing these technologies.
- Kieran Hawkins: The languages I'm most proficient in are Python and C++. I understand deep learning and neural networks and have implemented them from scratch in Python and C++, and I also have a decent understanding and some experience with PyTorch. I've also trained models using cloud-based GPUs on Lambda Labs and have experience with deploying and inferencing with the Llama 3 model on AWS, GCP, and Lambda Labs.

Questions:

Q: Would you recommend any other semantic segmentation model to be used for this application?

A: He recommended and provided us with a repo to one of the models on AI Hub that's tuned specifically to food. He gave us the option to try object detection or classification rather than semantic segmentation, as we will have an easier time with it due to the time we have.