Capstone Weekly Report

During this week's progress on my capstone project, I encountered some challenges with our initial plan. Originally, I had intended to collect food images from Instagram using their API. However, Instagram has recently implemented stringent regulations on API usage, primarily to safeguard user privacy. Unfortunately, these new restrictions meant that I couldn't access Instagram consumer accounts unless I was working for Business or Creator Instagram accounts. Moreover, collecting tens of thousands of food images, each posted by individual users, required sending queries one by one, which was impractical given me limited time frame.

To overcome these obstacles, I embarked on a quest to find alternative sources of data. Fortunately, I was successful in identifying two valuable datasets online that would serve as excellent substitutes. One dataset was obtained from Kaggle, while the other was sourced from the University of Barcelona. Combined, these datasets encompassed a total of 127 distinct food classes and a staggering 28,699 images, offering me a rich and diverse pool of data to work with.

In preparation for the next phase of my project, which involves data preprocessing, I initiated some essential groundwork. Firstly, I took the necessary steps to secure a free AWS (Amazon Web Services) account, a crucial resource for managing and processing the voluminous image data. I diligently studied the relevant documentation to understand how to seamlessly integrate Python with AWS. This involved learning how to launch an instance from the AWS platform and configuring SSH settings to ensure a smooth connection between my local environment and AWS.

In the upcoming week, my primary focus will shift towards data preprocessing. This pivotal step includes a myriad of tasks such as data augmentation, resizing, normalization, data splitting, data loading, label encoding, addressing class imbalance, shuffling data, and data visualization. Each of these operations is essential to ensure the quality and usability of our dataset, ultimately paving the way for robust and accurate model training.

As I move forward, I remain committed to addressing any challenges that may arise and continuously refining our approach to achieve our project's objectives. With the foundation laid during this week, I was well-prepared to dive into the intricacies of data preprocessing and continue making progress towards my ultimate goal of building a successful image classification model.