Final Project Proposal: Analyzing Perceived Identity in Gig Worker Profile Image Using Large-Scale Computing

Introduction

In this project, I aim to analyze the impact of perceived racial and gender attributes on the professional visibility of gig workers on platforms such as Fiverr and Upwork. Utilizing the FairFace dataset and Dask, I will process and analyze over 24,000 gig worker profiles, with a goal to expand the dataset to include a broader demographic across multiple platforms.

Relevance to Social Science

This study examines how perceived identity attributes affect professional interactions and opportunities within the digital gig economy. Through facial recognition technology, I aim to provide insights into discrimination and bias in both AI-based and non-AI-based environments, thereby enriching the discourse on social dynamics within gig work platforms.

Data Collection and Expansion

My current dataset includes over 24,000 profiles. To deepen my understanding of the gig economy, I plan to expand this dataset from platforms like Fiverr, which hosts over 380,000 active sellers, and Upwork, with its 12 million freelancers. As these platforms do not provide demographic data, it would be helpful to use large-scale computing to do facial recognition to infer racial and gender attributes.

Scalable Data Processing with Dask

I will use Accelerating Dask for efficient and scalable data processing. Dask allows for the parallel processing of large data volumes by dividing tasks into smaller chunks, which enhances performance and adapts dynamically to the size of the dataset and available computing resources.

Research Schedule

May 6-8: Data preprocessing and initial model testing.
May 9-16: Implement and refine parallel computing with Dask.
May 17-24: Conclude data analysis, finalize GitHub documentation, and prepare the project presentation.