AAI/CPE/EE-695A: Project Group 3 Team Contract

Michael Dasaro, Christopher Morales, Hamza Mustafa, Abdul Naeem Shaik

Problem statement:

Hunger is a serious problem worldwide. Technology has advanced incredibly far in recent years to the point where anyone with an internet connection has access to essentially all of human knowledge at their fingertips. With the advent of uber eats, grubhub and the like, people can order food to be delivered straight to their door, even sometimes being delivered by a robot. As such it is disconcerting to learn that roughly 10% of the world's population - that's 830 million people - go to sleep hungry at night. Despite this, the issue is not one of scarcity but of distribution. Worldwide an estimated 17% of all food production is wasted, which adds up to an astronomical 1.3 billion tons of food wasted in a year. This much food can go a very long way to helping those hungry people worldwide. Our goal with this project is to analyze data for food scarcity and food surplus worldwide in order to determine an optimal solution for food distribution to those who need it most. Perhaps as an extension to the scope of this project, we can look into methods of distributing such food.

Description of data:

Food surplus information:

- https://www.kaggle.com/datasets/joebeachcapital/food-waste
    - Food waste per country
- https://www.kaggle.com/datasets/yanchoo/global-food-waste-2000-2021
    - Food waste per country over time
- https://www.kaggle.com/datasets/sennyben/food-waste-2021
    - Food waste, more geographically detailed

Food scarcity information:

- https://www.kaggle.com/datasets/maryamsikander/sdg-2-zero-hunger
    - Not a perfect dataset but can be used along population stats to get the number of undernourished and number of severe food scarcity per country

Shipping cost information:

- https://www.kaggle.com/code/klmsathishkumar/shipping-cost-prediction
    - Not too pertinent but a good setup. We'll probably need an estimator for the cost of an entire container ship of food.

Implementation plan:

- Cleanup datasets into simpler bins of:
    - Countries that need food and the amount they need
    - Countries with food surplus and how much is wasted
    - A cost estimator for shipping between any two countries based on distance and amount of food
- Optimize for the most amount of food redistributed with the least cost:
    - Perhaps using a multifaceted decision tree approach where food distribution decisions are made based on the amount of hungry people and amount of food surplus in countries

Task allocation:

- Continue dataset search: *All group members*
- Cleanup datasets: *Hamza*
- Shipping cost estimator research: *Michael*
- Build shipping cost estimator: *Chris*
- Implement pipeline for training and testing with different models: *All group members*