Title:
Framework for Spatial Health Analytics Using Digital Exhaust, Crowd-Source Data, and Electronic Health Records

Abstract:
Low accessibility to healthy and nutritious food has been hypothesized to increase health disparities. Low-accessibility areas, called food deserts, are particularly commonplace in lower-income neighborhoods. However, indices for modeling food desert intensity are subjectively defined, and there is little agreement in the literature on their validity or relative strength. Moreover, such indices are largely based on census data, which limits their frequency and geographical resolution to that of the census.
In this dissertation, we first propose an assessment framework for objectively defining and comparing the utility of food desert indices using machine learning models. We introduce the concept of food desert index utility score, based on which we can compare the strength of indices for describing the food environment. We then focus on the Metro Atlanta area in Georgia, USA, as a case study to explore the effect of the geographic spatial resolution of models and the impact of adding or ignoring neighborhood-level income or vehicle access on the utility of food desert indices.
We then use demographic, geographic, and health data, as well as real-time data from platforms such as Yelp, Google Maps and crowd-sourced data using Amazon Mechanical Turks, to build a food desert index that has both higher spatial and temporal resolution than the standard Food Desert measures currently used in the US Food Reseach Atlas Data Base (FRADB). The new food desert index can be used, for example, to measure a person's exposure to varying food environments during a commute, to analyze the effect of the food environments on health outcomes or nutrition behaviors, or to suggest behavioral changes to reduce exposures. We use this temporally and spatially high-resolution, context-aware index (factoring in pseudo-real-time traffic density) in a concept application that suggests alternative routes with similar ETAs between a source and destination in the Atlanta metropolitan area to expose a traveler to better food environments. The resulting model was sensitive to changes in the environment that occurred after the census data was collected. In addition to informing community planners and policymakers more accurately than traditional food desert indices in the FRADB , our novel food desert index allows us to measure environmental effects on individuals and suggest personal behavioral changes, this is showcased in the fact that our index was sensitive to a new healthy food outlet being opened after the FRADB data were collected and suggested routes that were in the proximity of this newly opened center. In particular, an evaluation of 248,000 routes from random locations to 28,000 food retailers demonstrates that the Atlanta food environment creates a strong bias towards eating out rather than preparing a meal at home when access to vehicles is limited.
Finally, as a case study, we analyze the hospitalization data of more than 64,000 COVID-19 patients in the Metro Atlanta Area in the first 19 months of the pandemic and observe that living in a food desert was associated with a higher per capita number of patient hospitalizations and higher number of deaths per hospitalizations in the first months of the pandemic.

Committee:
Prof. Gari D. Clifford
Prof. Matt Reyna
Prof. Reza Samani