Team 0

“Microbiome Data Bank”

Anujit Sarkar, Ph.D.
Charley Wang, Ph.D.
Elizabeth Miller, Ph.D.
Guy Dayhoff & Mitch Hastings

Motivation

- Can unsupervised machine learning detect geographic similarities in microbiome data?
- Historically, human bias has impacted how human biological variation is classified. Can machine learning detect meaningful patterns of variation without including human bias?
- Can a large-scale data bank of microbiome data be established for global usage and assist in answering geographic related research questions?

Execution

- Web App (Django)
- Regional Microbiome Data (16S)
- Unsupervised Machine Learning (k-means/PAM)

Outcomes

- Users can contribute Microbiome data to the MDB
- Users can download the pipeline for local usage
- Users can select regions for cluster comparison
- Intuitive design for non-technical users

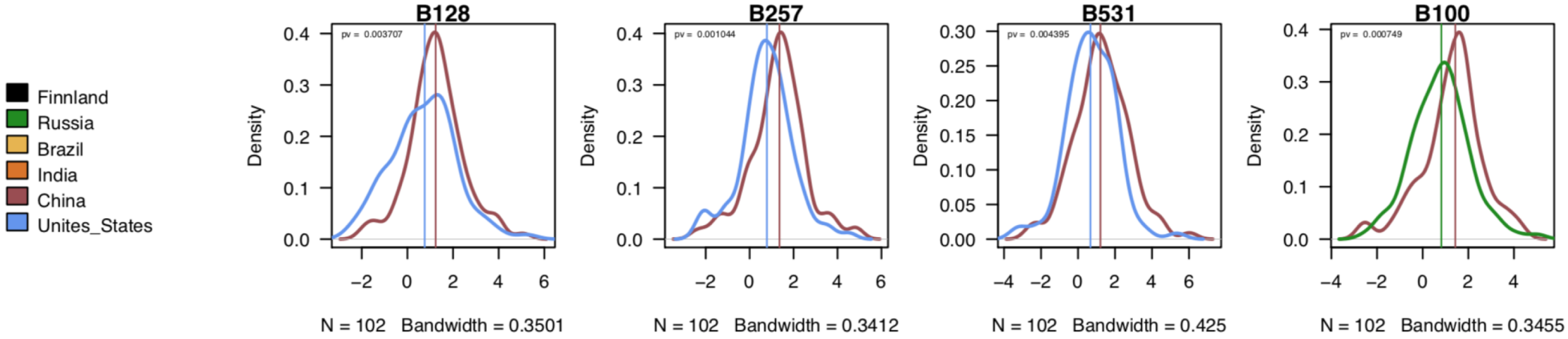
Day 1 Progress Report

- AWS Instance Up and Running

- Data Preprocessing
  - Mined Data from previous published lit’
  - Pulled 4175 samples
    - All samples are from healthy adults
    - All samples the V4 region of the 16S RNA gene has been studied
    - 7 Regions Represented:
      - China (1995)
      - Russia + Finland + Estonia (1000)
      - Brazil (203)
      - India (78)
      - USA (899)

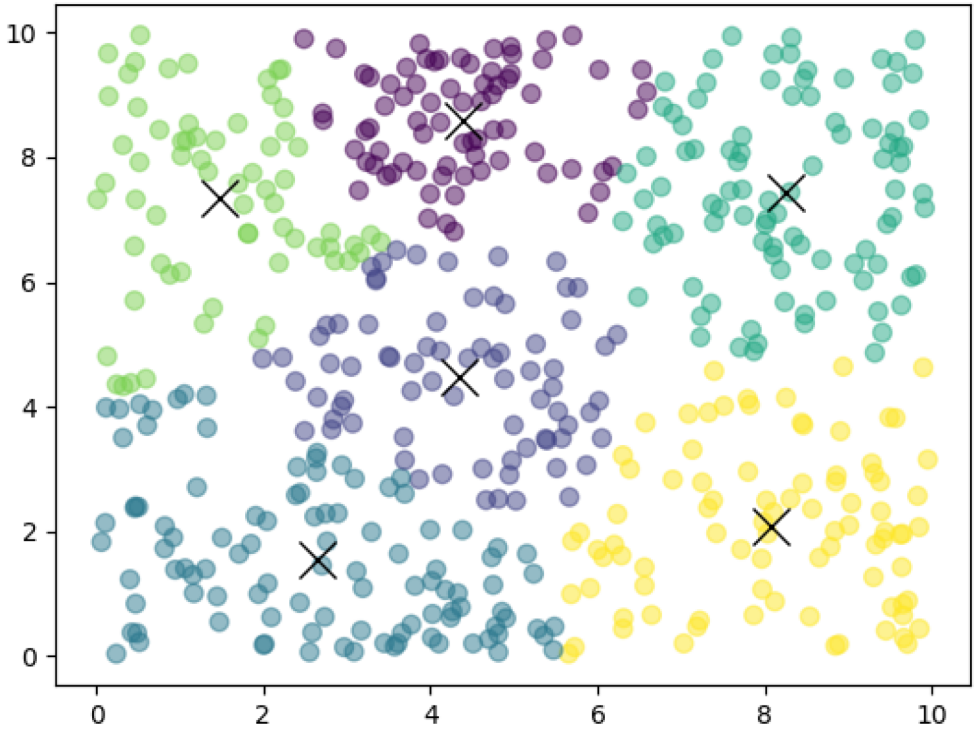
Day 1 Progress Report (cont.)

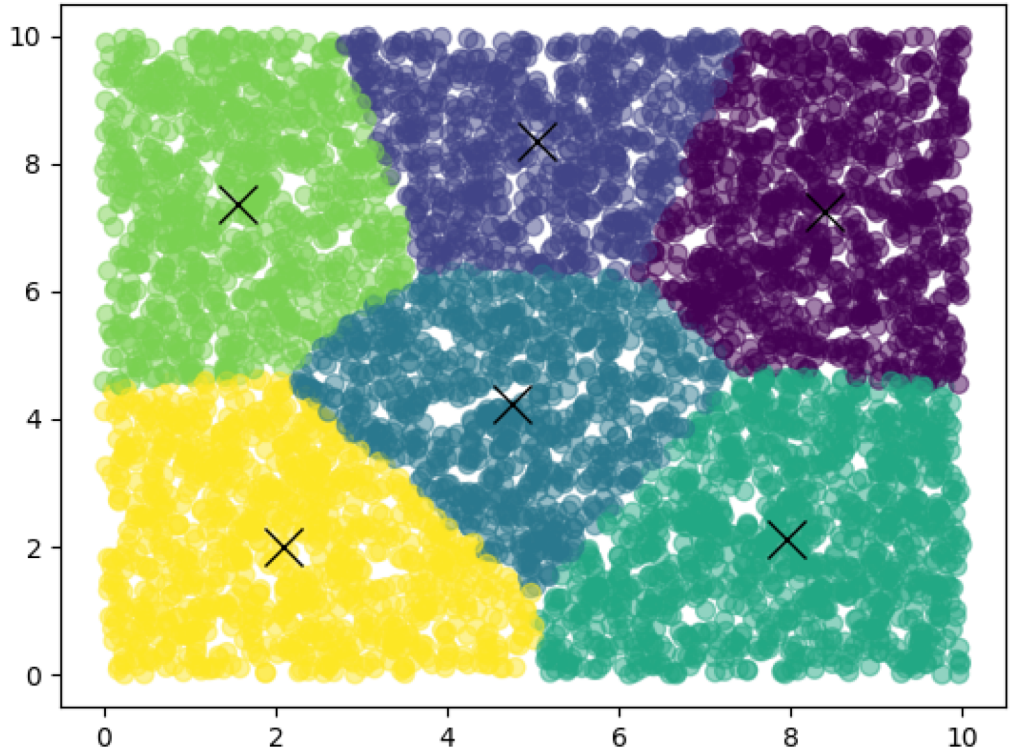
- Feature Selection to inform ML



Day 1 Progress Report (cont.)

- K-means clustering implemented in TF





Day 2 Progress Report

Task	Status
Data Acquisition	Yes
Data Preparation	No
Data Mining	In Progress
Web Server Setup	Yes
Web App Setup	No
UI Design	No
UI Implementation	No
Write Up	No
Model Implementation	In Progress

Day 2 Progress Report (cont.)

Task	Status
Data Acquisition	Yes
Data Preparation	Yes
Data Mining	In Progress
Web Server Setup	Yes
Web App Setup	Yes
UI Design	Yes
UI Implementation	In Progress
Write Up	In Progress
Model Implementation	In Progress

Day 2 Progress Report (cont.)

- Data Preparation
  - Transitioned to subsampling (190GB=> 93GB => 5.2GB)
- Web App Setup
  - Django web-app live
- UI Design & Implementation
  - Simple, Easy to Understand, and Intuitive