



CS5100 Final Project

Jeff Turgeon
Xichen Liu

Final project will be applying Agglomerative Hierarchical Clustering algorithm to a selected dataset and using various heuristic functions to determine the actions of agent.



Agglomerative Hierarchical Clustering

- Clustering from every single data.
- Performs good on messy dataset in unsupervised learning
- Agent is the program
- Actions are the combinations made to cluster the data
- Different heuristic functions to determine the actions



Distance Metrics - Affinity

We will compare the following distance metrics to determine which one is most optimal for our given dataset:

- Euclidean Distance
- Manhattan Distance
- Cosine Distance
- At least 2 Heuristic Functions based on the chosen dataset