

Special Topics: Data Analytics and Visualization in Healthcare
CSCI-GA.3033-096 (19635)
Lab assignment 6

Instructions:

Part I

Using the dataset `Wisconsin_breast_cancer.csv`, perform the following exercise:

- Investigate how to implement Hierarchical Agglomerative Clustering in either R or Python.
- Create a dendrogram.
- Evaluate your model.

Part II

- Select a current (no more than 5 years), journal/conference article pertaining to machine learning in healthcare.
- The paper should have at least one unsupervised machine learning method.
- Summarize the key ideas discussed in the paper.
- Critique the overall content of the paper.
- Include a reference to the paper in your write-up using a citation format.

Notes:

- This assignment is individual.
- You need to submit 2 files (the .R or .ipynb file and the MS Word/PDF document).
- Please do not copy/paste information. Plagiarism will be penalized.