Galaxy Morphology Classification using Unsupervised Machine Learning

Presented by: James Liu

Imaging Science, PhD

Advisor: Jeyhan Kartaltepe

Overview

- Motivation
- Background
- Methodology
- Results
- Future Work & Challenges

Motivation

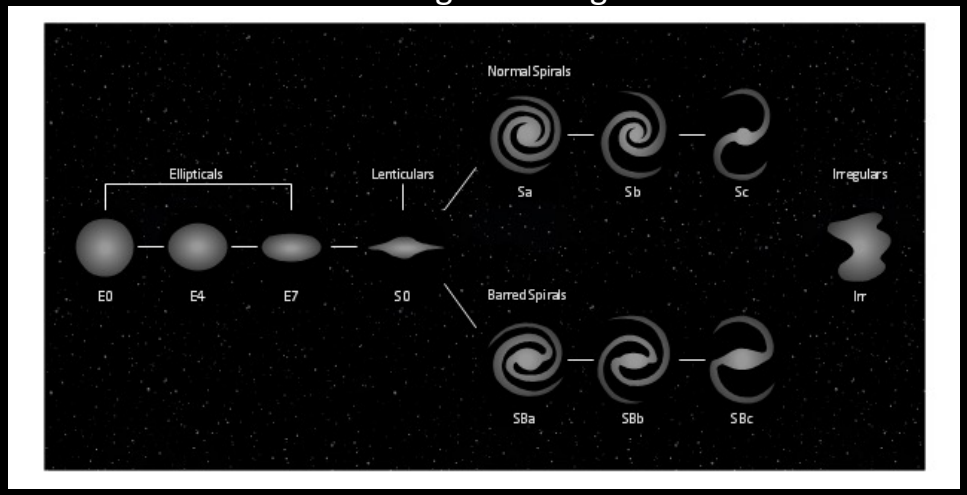
- Understanding of Galaxy Evolution
  - Morphology Classes provides insight into Galaxy Structure, formation & evolution
- Overabundance of data from new state-of-the-art telescopes like James Webb Space Telescope (JWST)
  - Classification of Galaxy was previously conducted by only a handful of expert astronomers
  - Reliance on Citizen Science Projects like Galaxy Zoo for classification

Objective: Classify Galaxies in JWST Images using Unsupervised Machine Learning Method

Background

- Galaxy morphological classification
  - Categorize galaxies into groups based on their visual appearance
  - The Hubble Sequence invented by Edwin Hubble in 1936, is a morphological classification scheme for galaxies

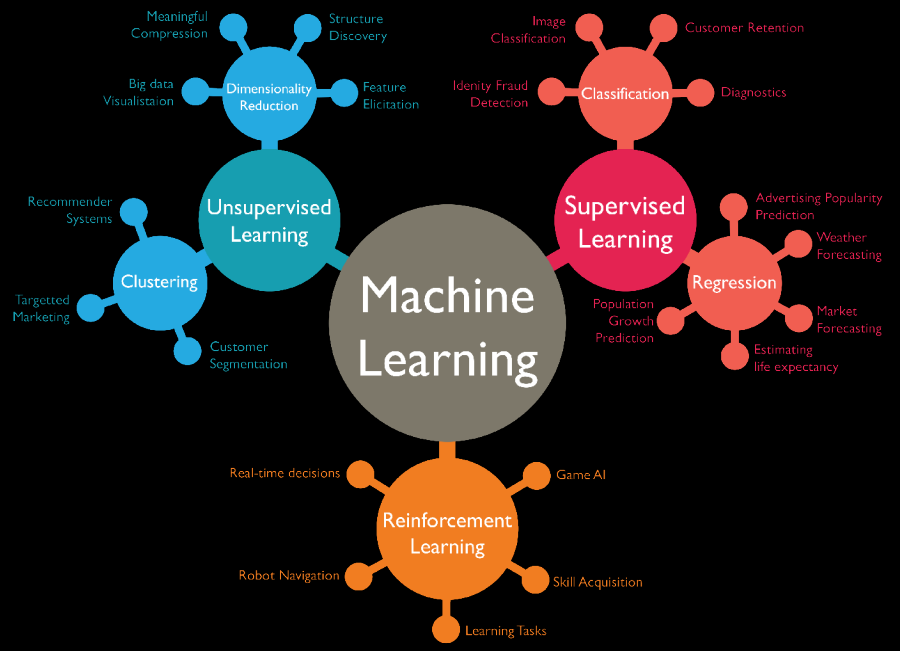
Hubble Tuning Fork Diagram



Ref:
https://www.conceptdraw.com/examples/fork

Background

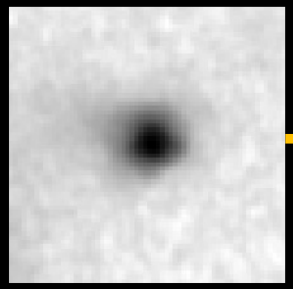
Ref: https://lawtomated.com/supervised-vs-unsupervised-learning-which-is-better/



Ref: https://towardsdatascience.com/machine-learning-types-2-c1291d4f04b1

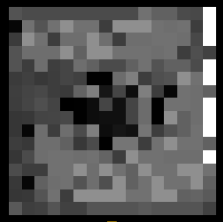
Background

Input Images
(64x64)

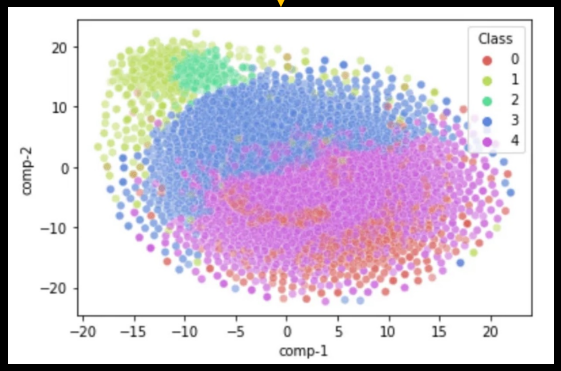


Encoder

Latent Space
(16 x 16)

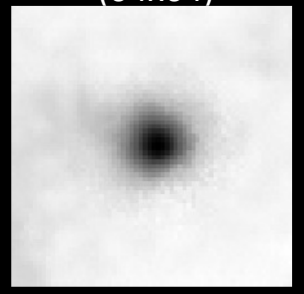


Clustering



Decoder

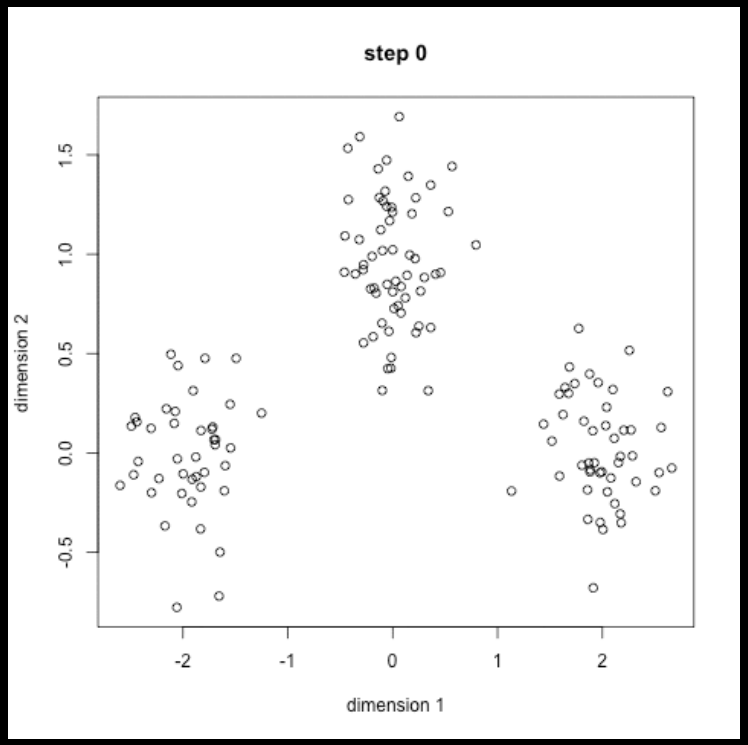
Reconstructed Images
(64x64)



Background

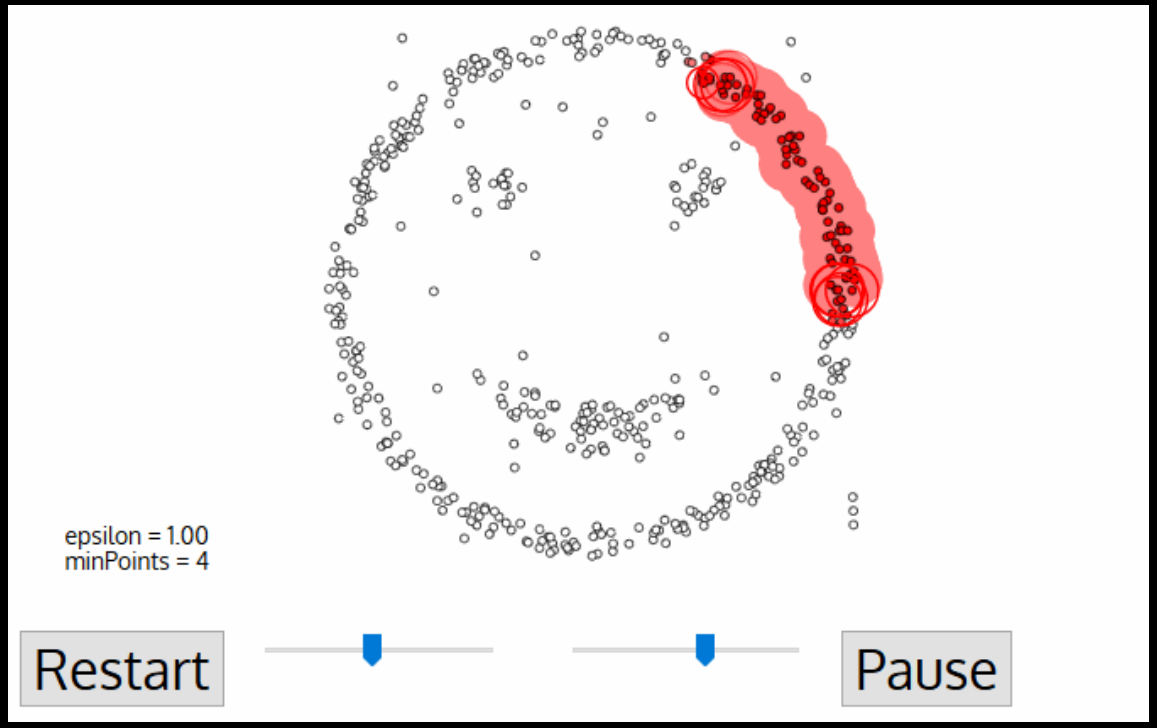
- Centroid-Based Clustering
  - K-Means
- Density-Based Clustering
  - DBSCAN (Density-Based Spatial Clustering of applications with noise)
- Distribution-Based Clustering
  - Gaussian Mixture Models
- Hierarchical-Based Clustering
  - Agglomerative Hierarchical Clustering

Centroid-Based Clustering (K-Means)



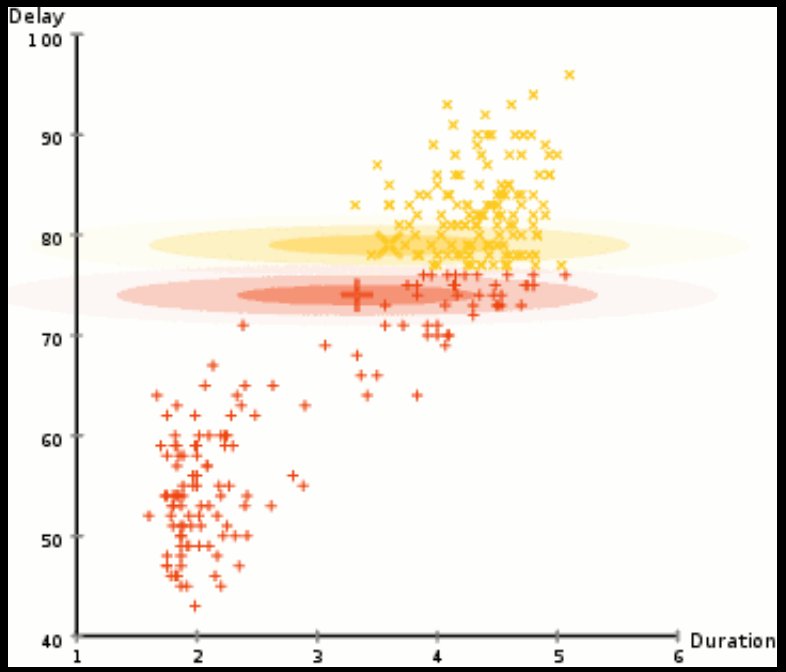
Ref: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

Density-Based Clustering



Ref: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

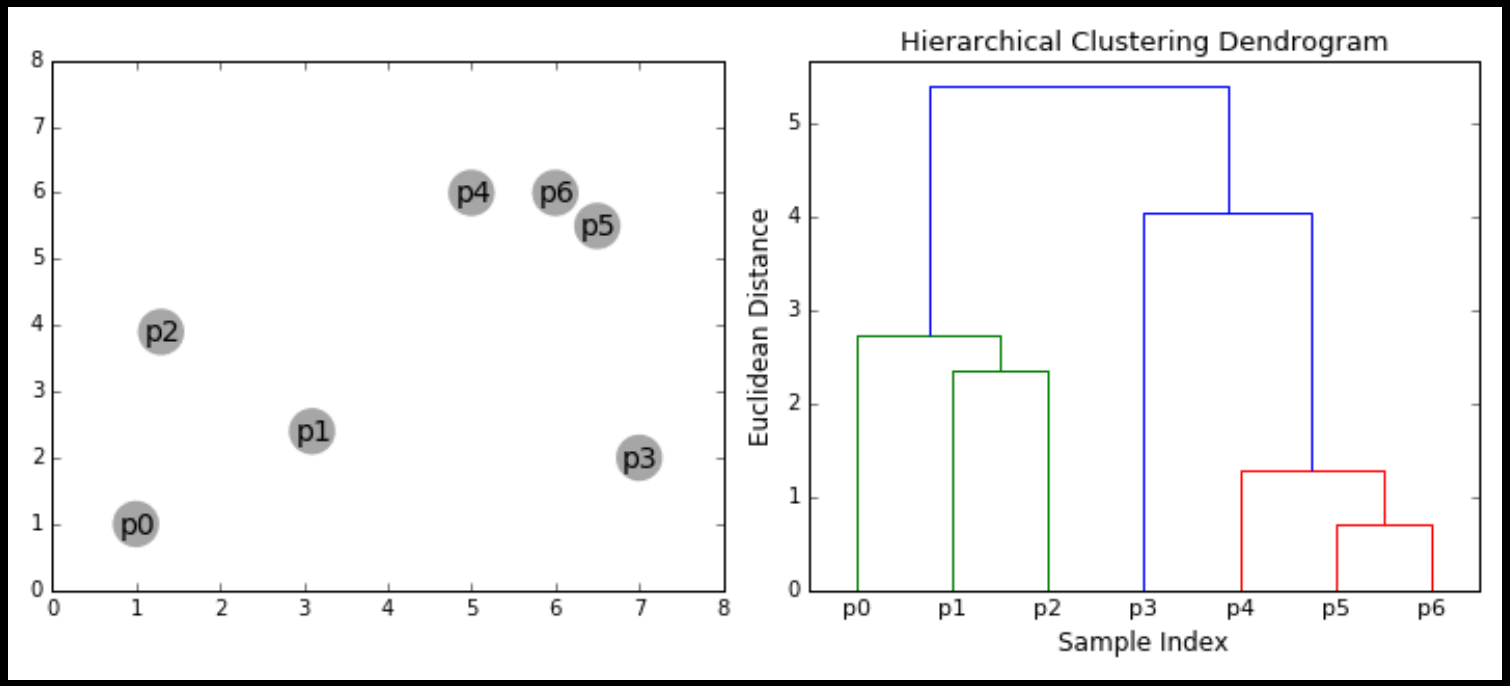
Distribution-Based Clustering



Ref: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

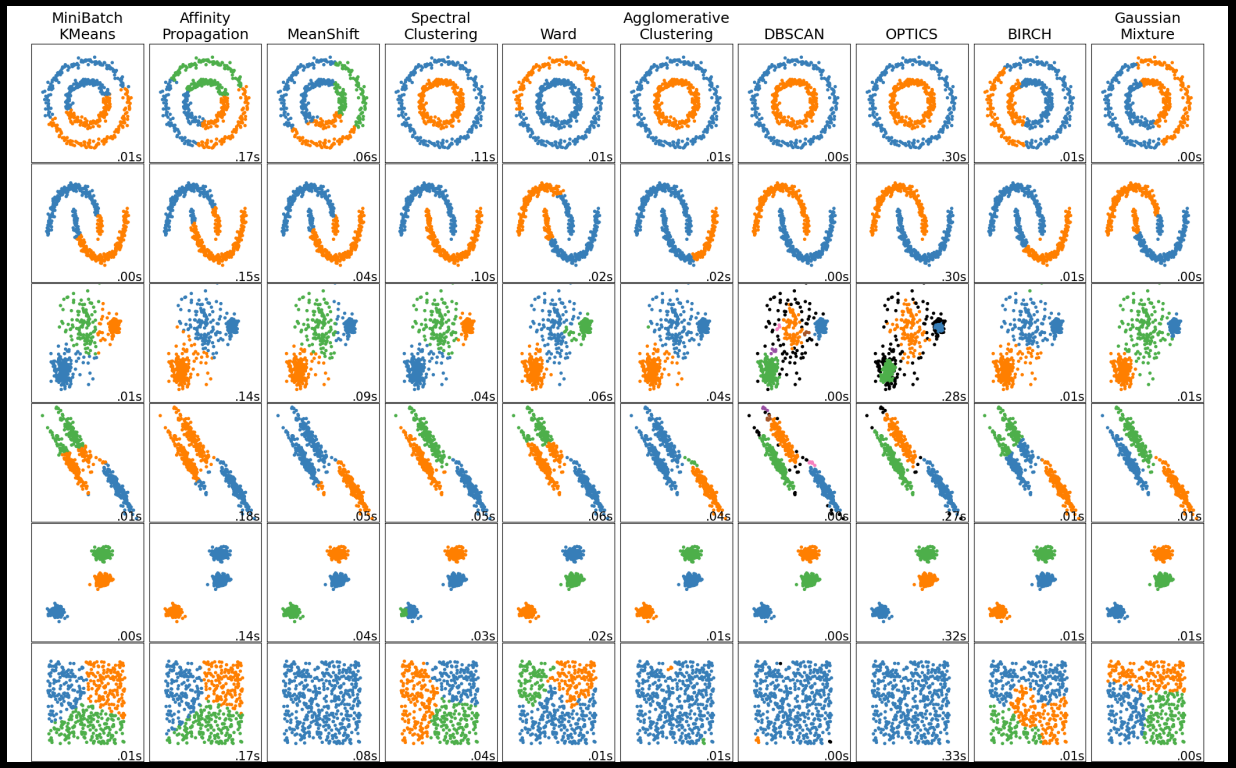
Hierarchical-Based Clustering

Agglomerative Hierarchical Clustering



Ref: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

Overview of Clustering Results



Ref: https://scikit-learn.org/stable/modules/clustering.html

Methodology

- Simulated JWST mosaics (large surveys of the sky) are used to create stamps (individual cutouts of galaxies)
  - IllustrisTNG100 with added noise to simulate images taken from JWST
  - Correspond to those taken via JWST's NIRCam between wavelength range of 1-4 microns
    - Filters: F115W, F150W, F200W, F277W, F356W, F444W
- A total of 10602 stamps were generated (1767 stamps created in each Filter)
  - Format: FITS (Flexible Image Transport System)
  - Image Size: 64 x 64 pixels
- Vector Quantized Variable Autoencoder (VQ-VAE) is used to reduce the dimensionality of the stamp data into a discrete latent representation
- Clustering of the lower dimension image data is done using a clustering algorithm.
  - K-means

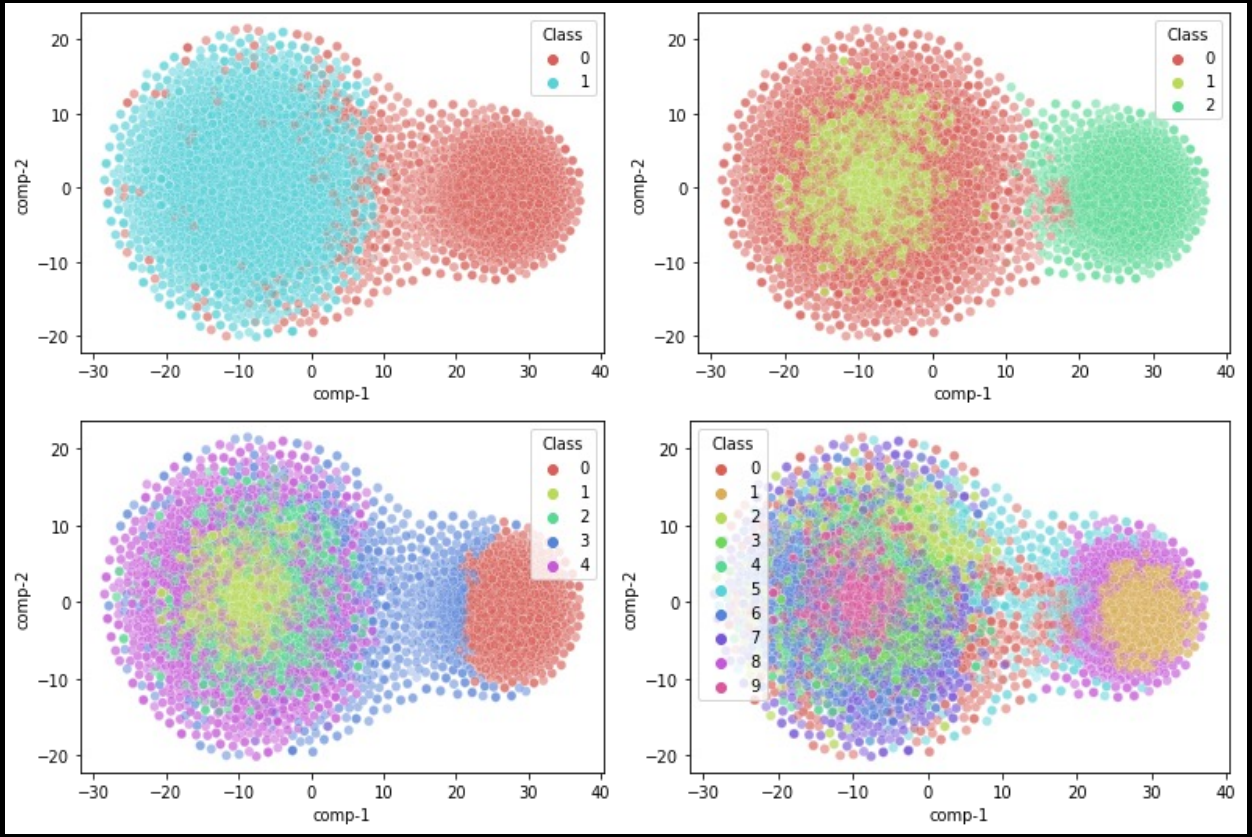
Results

Original

Latent Space

Reconstruction

Results


Class
0
1
comp-1
comp-2
Class
0
1
2
comp-1
comp-2
Class
0
1
2
3
4
comp-1
comp-2
Class
0
1
2
3
4
5
6
7
8
9
comp-1
comp-2

Future Work & Challenges

- Because JWST is still in its infancy, there is very little labeled morphology data to validate prediction/classification accuracy
  - Other methods to evaluate clustering results currently being explored
- Reliability of K-means on this dataset
  - More research into K-means should be explored to understand impact on high-dimensional data
  - Testing results from other types of algorithms might provide different insight
  - Try different clustering algorithm/methodology that might provide different results i.e Hierarchical Clustering, HDBSCAN, Spectral Clustering, etc.

Summary

- Motivation
- Background
- Methodology
- Results
- Future Work & Challenges

Questions?

Quiz Time!

- Go to mycourses to answer the question