



k-means clustering

Due dates for this week:
12pm Friday — practicum 9

Kmeans.PY

Next week:

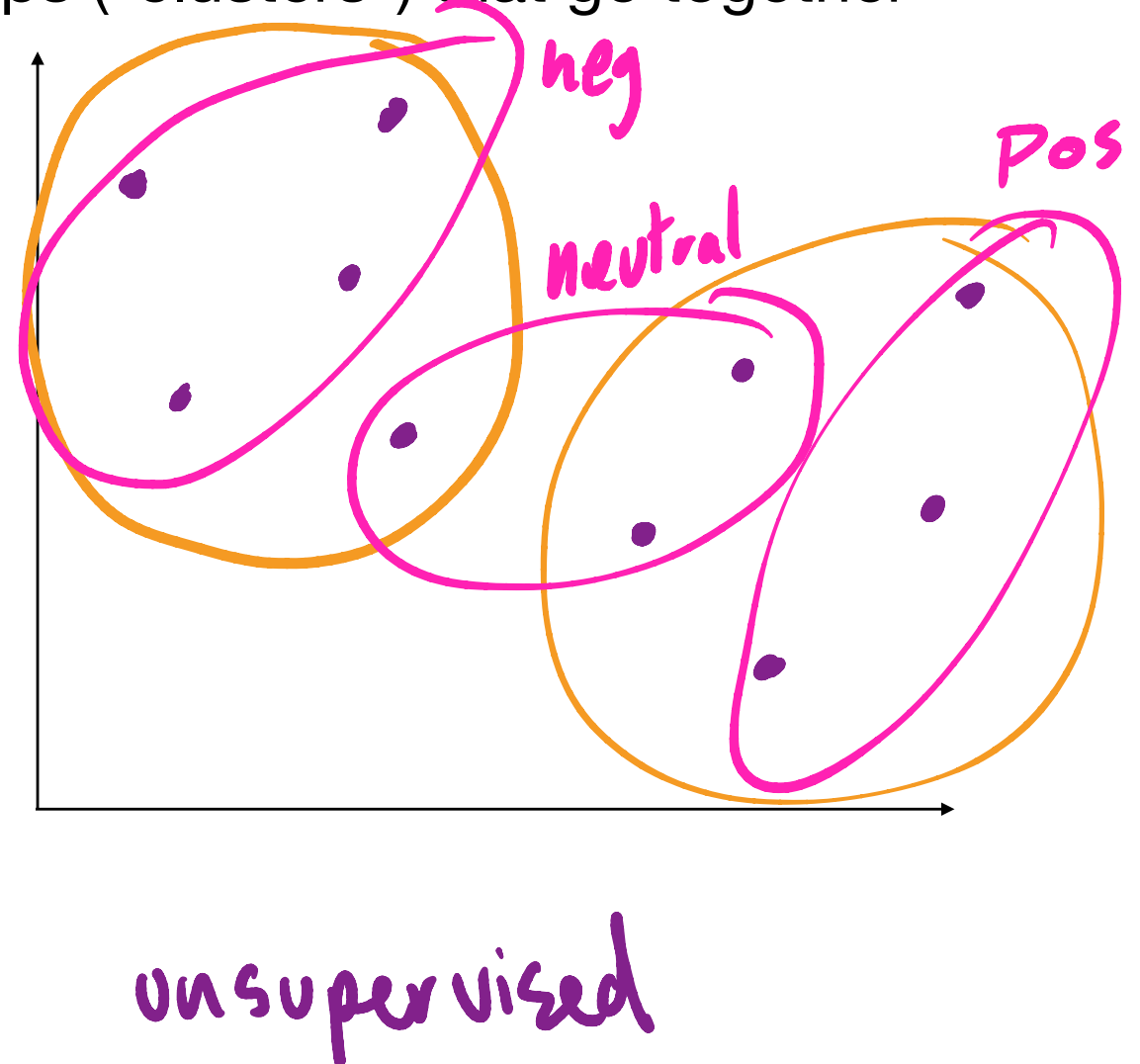
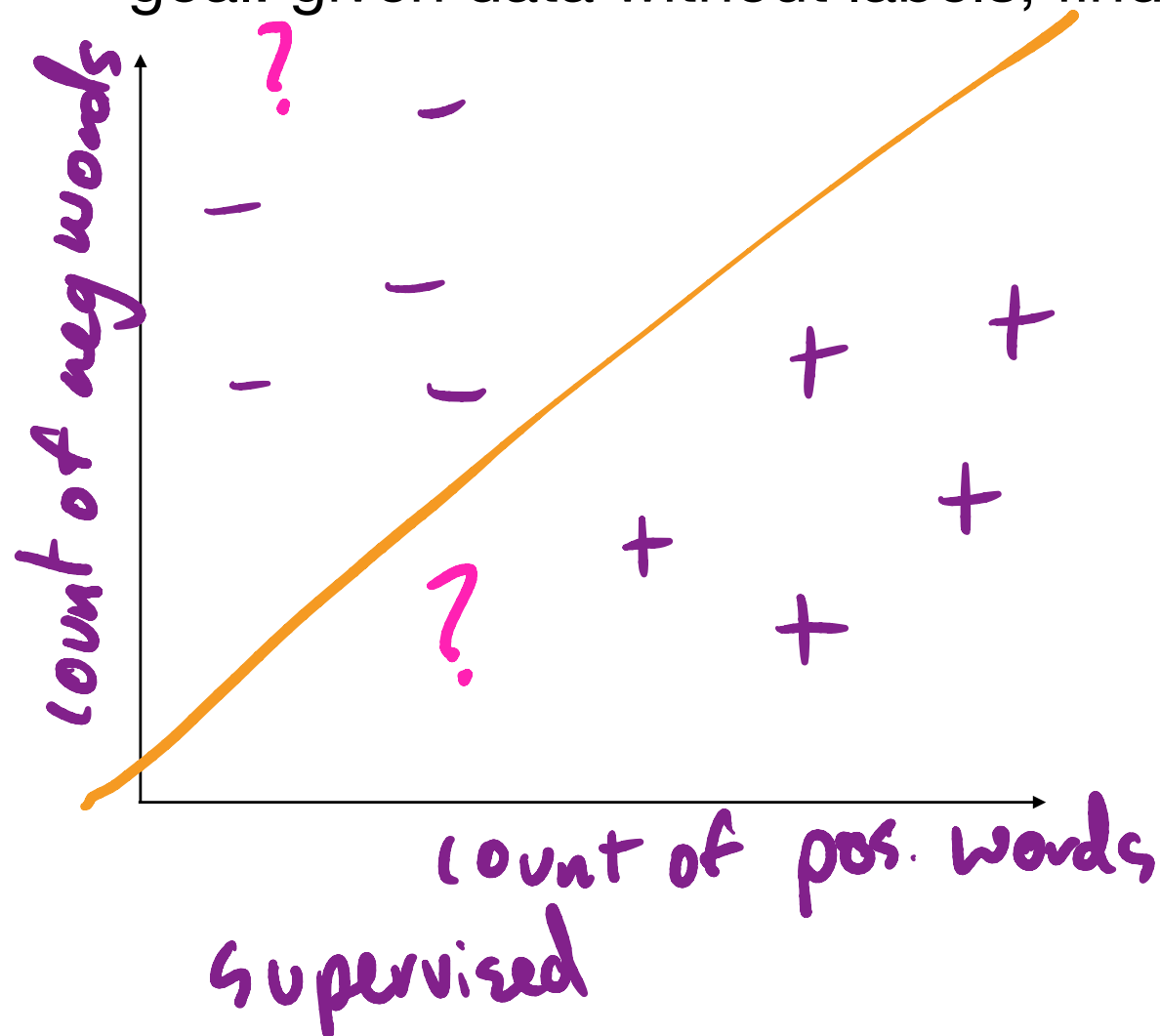
Office hours/in-person help time for **all final projects** during regularly scheduled Wednesday sections.

No class on 11/11 (during Thursday sections), it's Veteran's Day

All of Felix's office hours through thanksgiving will be on Calendly

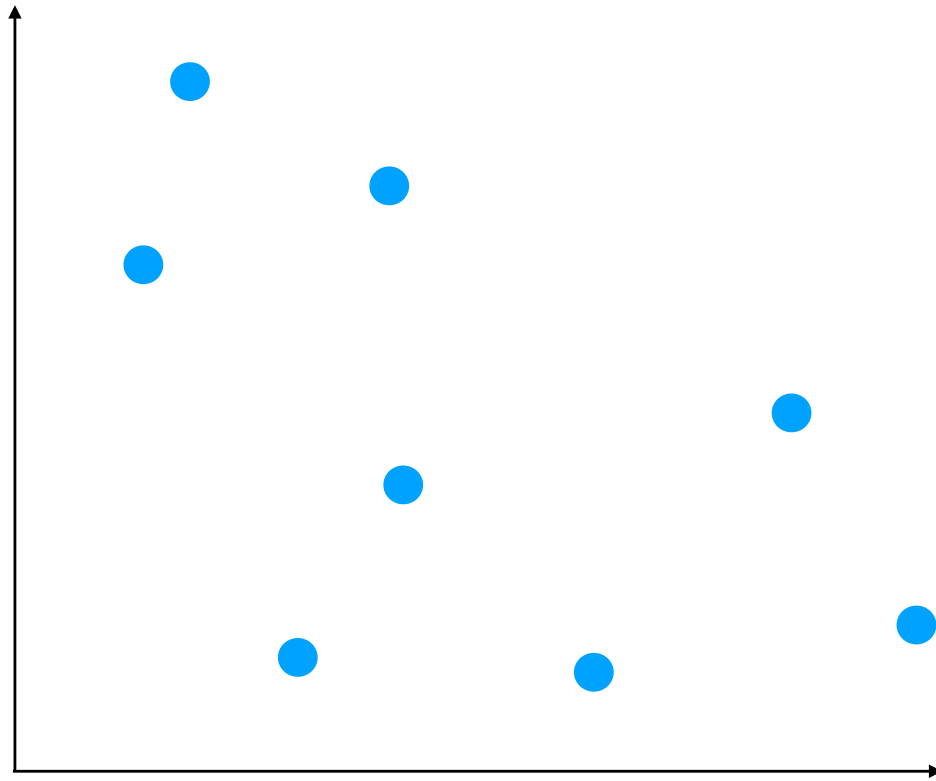
k-means clustering

- unsupervised machine learning algorithm
- goal: given data without labels, find groups ("clusters") that go together



k-means clustering

- Step 1: plot your data

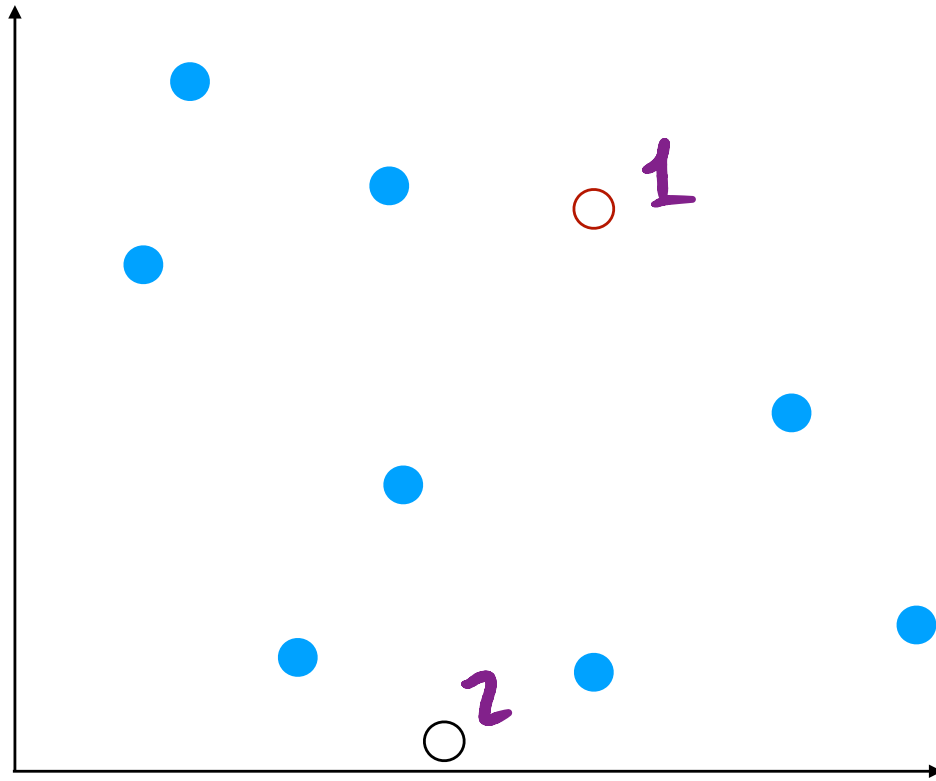


new type: "tuple"

k-means clustering

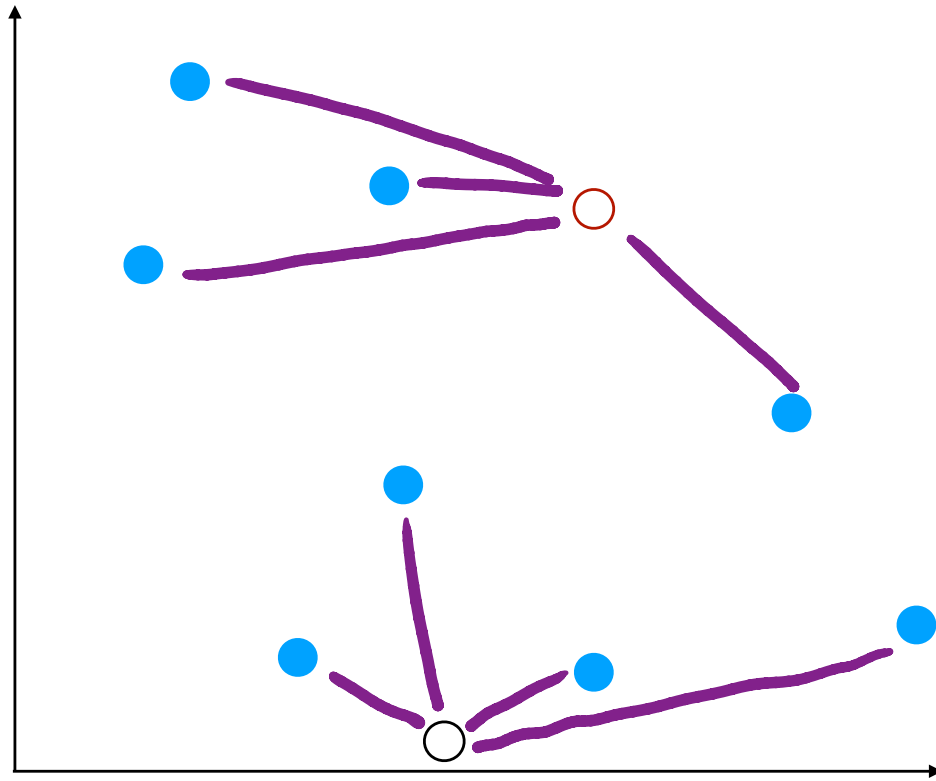
→ 2 in this case

- Step 2: randomly initialize k means ("centroids")



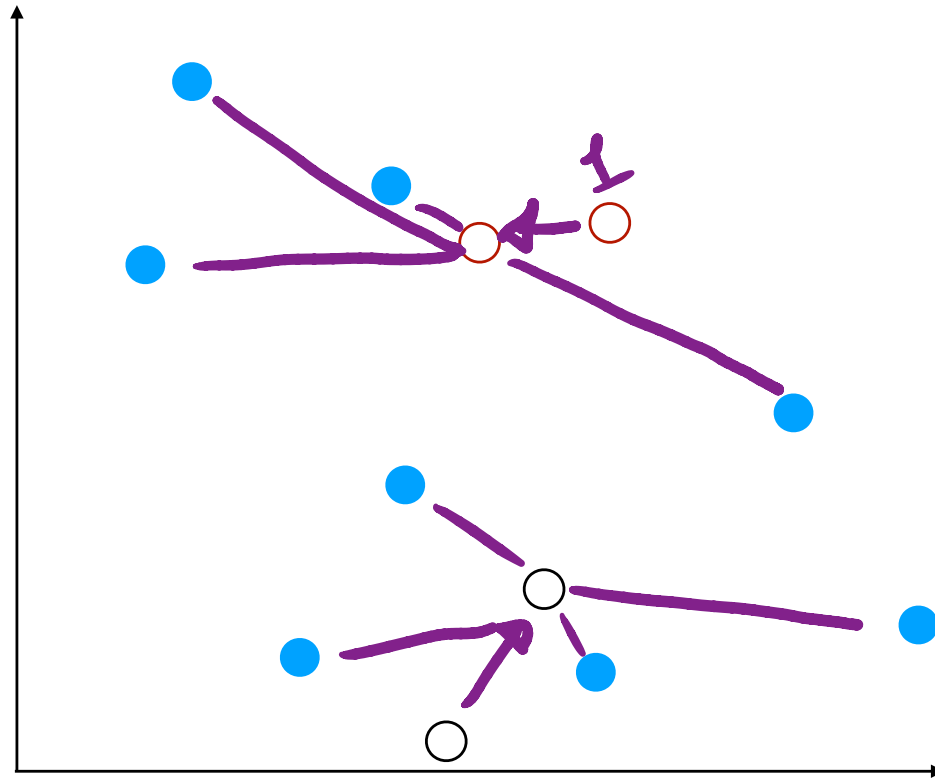
k-means clustering

- Step 3: assign data points to closest centroid



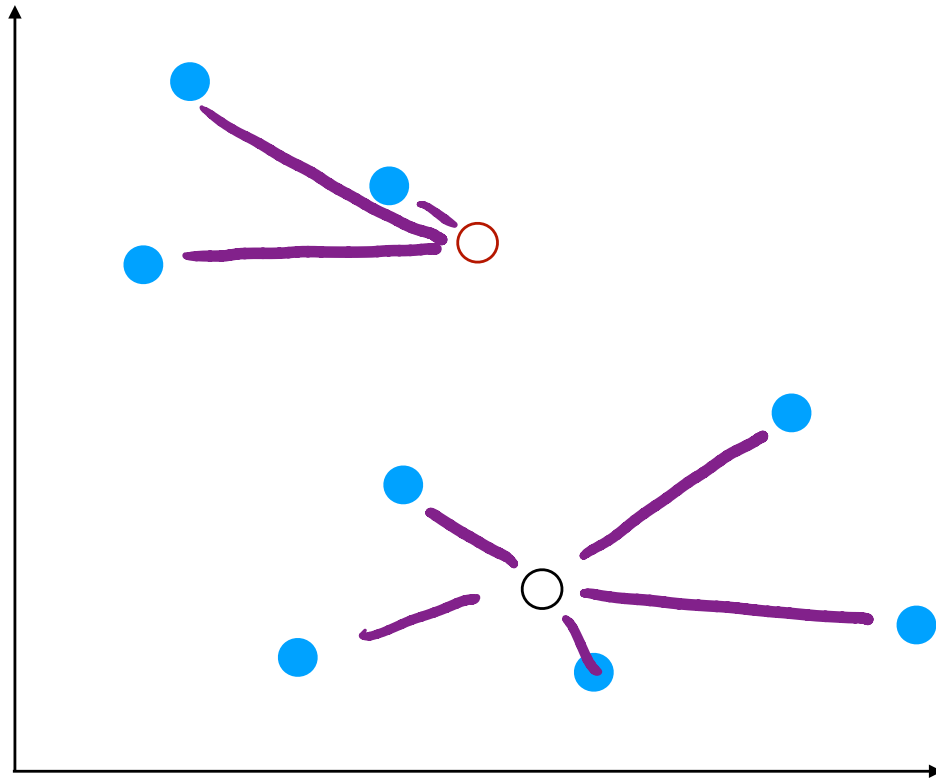
k-means clustering

- Step 4: adjust the centroids to be the mean of the group of data that they "belong" to



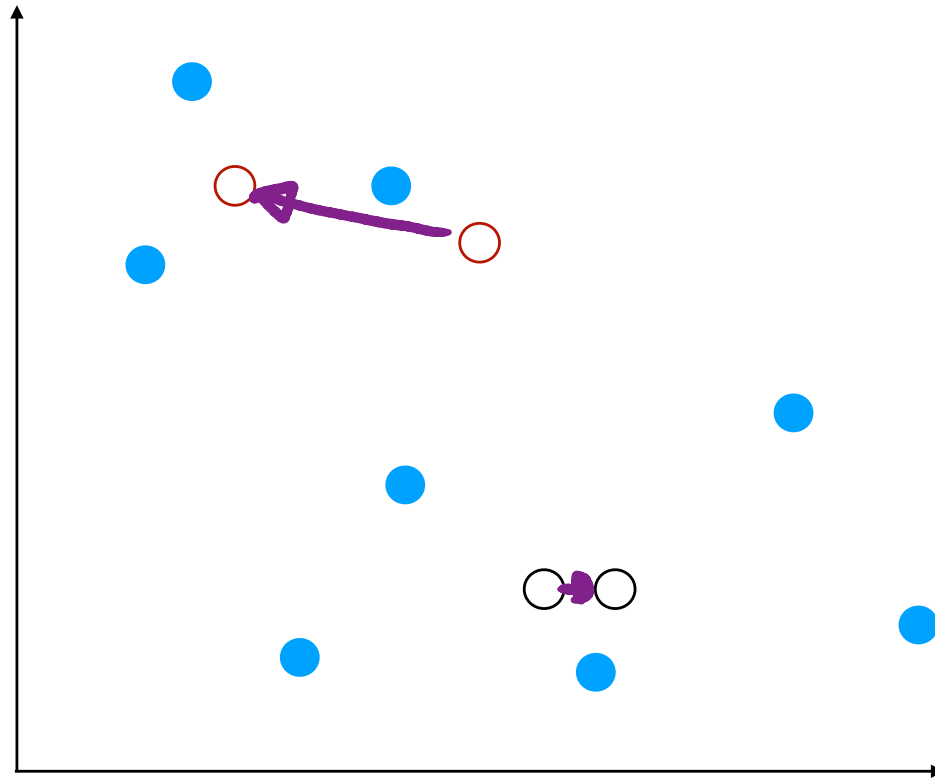
k-means clustering

- Step 3: assign data points to closest centroid



k-means clustering

- Step 4: adjust the centroids to be the mean of the group of data that they "belong" to



k-means clustering

- Terminate either after a pre-determined number of iterations or when centroids no longer move *→ no points change cluster membership*

