DS-630 Discussion Week 9

What is label propagation? Why would you implement it, and how?

Label propagation is an algorithm for finding communities. In comparison with other algorithms label propagation has advantages in its running time and the amount of a priori information needed about the network structure (no parameter is required to be known beforehand).

The label Propagation Algorithm is one of the semi-supervised machine learning algorithms that assigns labels to unlabelled data observations in order to partition the classification of data observations within the dataset.

Each of the distinct nodes has its corresponding label

The corresponding label of each node denotes the distinct community that this node belongs to

Through iteration within the network, each node will update its belongingness community based on the belongingness community of neighbor nodes until all of the nodes with the same label are attributed to the same communities

The updated community of each node will be the corresponding belongingness community of the maximum number of nodes

Eventually, densely connected nodes reach a common-label community

What is the difference between anomaly detection and novelty detection?

Data that is in distribution can be called novelty data. Novelty detection is when you have new data (i.e. OOD) and you want to know whether or not it is in distribution. You want to know if it looks like the data you trained on. Anomaly detection is when you test your data to see if it is different than what you trained the model. Out-of-distribution detection is essentially running your model on OOD data. So one takes OOD data and does novelty detection or anomaly detection (aka outlier detection). Novelty detection is a statistical method used to determine new or unknown data and determine if these new data are within the norm (inlier) or outside of it (outlier). Anomaly detection is the process of finding outliers in a given dataset. Outliers are the data objects that stand out amongst other objects in the dataset and do not conform to the normal behavior in a dataset.

When a person engages in negative self-talk, they can perceive themselves as failures, leading to a self-fulfilling prophecy. Or, they can choose to focus on an ideal to which they aspire such that it comes to represent reality to them, whether or not it matches how others perceive them. Once the modification has been accepted, you can create a model practice for ethics that places an emphasis on objectivity, integrity, and causing the least amount of harm possible.

Clustering the images using kmeans