HW4

1. $dist = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$

$Obs_1 = \sqrt{(0-0)^2 + (3-0)^2 + (0-0)^2} = \sqrt{9} = 3$

$Obs_2 = \sqrt{(2-0)^2 + (0-0)^2 + (0-0)^2} = \sqrt{4} = 2$

$Obs_3 = \sqrt{(0-0)^2 + (1-0)^2 + (3-0)^2} = \sqrt{1+9} = \sqrt{10}$

$Obs_4 = \sqrt{(0-0)^2 + (1-0)^2 + (2-0)^2} = \sqrt{1+4} = \sqrt{5}$

$Obs_5 = \sqrt{(-1-0)^2 + (0-0)^2 + (1-0)^2} = \sqrt{1+1} = \sqrt{2}$

$Obs_6 = \sqrt{(1-0)^2 + (1-0)^2 + (1-0)^2} = \sqrt{1+1+1} = \sqrt{3}$

2. Since $k=1$, select the nearest neighbor to $(0,0,0)$ which is Obs_5. ∴ the prediction is green.

3. Since $k=3$, select the three closest neighbors to $(0,0,0)$, which are Obs_5, Obs_6, and Obs_2. That is green, red, and red respectively. ∴ the prediction is red.

Each neighbor contributes equally to the prediction when using uniform weight.

When $k=3$ $P(Red) = 2/3$ and $P(Green) = 1/3$

All Obs $P(Red) = 4/6 = 2/3$ and $P(Green) = 2/6 = 1/3$

Distance weight:

$w_i = \frac{\exp(-d_i^2)}{\sum_{j=1}^{k} \exp(-d_j^2)}$ for $k=3$

Obs_5 green $w_5 = \frac{\exp(-1.41^2)}{\exp(-1.41^2) + \exp(-1.73^2) + \exp(-2.23^2)} \approx 0.413$

Obs_6 Red $w_6 = \frac{\exp(-1.73^2)}{\exp(-1.41^2) + \exp(-1.73^2) + \exp(-2.00^2)} \approx 0.288$

Obs_2 Red $w_2 = \frac{\exp(-2.00^2)}{\exp(-1.41^2) + \exp(-1.73^2) + \exp(-2.00^2)} \approx 0.299$

$P(Green) \approx 0.413$

$P(Red) \approx 0.288 + 0.299 = 0.587$

HW4 Continued

Steps of k-means algorithm:

1. Assign each data point to the nearest cluster centroid
2. Update the cluster centroid to be the mean of the data points assigned to it.

Proof:

Assign: Since the data points are being assigned to the nearest cluster centroid, the sum of squared distances between the data points and their cluster centroids. ∴ the objective function will stay the same or decrease.

Update: Since the cluster centroid is being updated to be the mean of its data points, the average distance between the points will decrease or stay the same.

Since both steps decrease the objective function or keep it the same, the objective function is non-increasing.

Steps of medians of data points w/ Manhattan distance

1. Initialize k centroids randomly
2. Assign data points to the nearest centroid based on Manhattan distance
3. Update centroids to be the medians of the data
4. Repeat steps 2 and 3 until convergence

Proof

Assign: Assigning the data points to the nearest centroid based on the Manhattan distance minimizes the sum of absolute distances from each data point to its centroid. ∴ the next iteration (J_{t+1}) will always be less than or equal to the current iteration (J_t)

Update: Updating the positions of the centroids to the medians of their assigned points ensures the centroids are positioned to minimized the sum of absolute distances. ∴ the next iteration (J_{t+1}) will always be less than or equal to the current iteration (J_t)

Since both steps decrease the objective function or keep it the same, the objective function is non-increasing.