

Assignment 5

Name: Soham Kamble
CWID: A20517098

Problem 1

1.1

$$Y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

[0]

[0]

[-1]

[-1]

[1]]

1.2

v1: No label

v2: label 2

v3: label 2

1.3

v1: label 2

v2: label 2

v3: label 2

1.4

v1: label 2

v2: label 2

v3: label 2

Convergence Formula:

$$F = [I - \alpha S]^{-1} * Y$$

Matrix After Convergence:

[[-0.0972985]

[-0.20919178]

[-0.20910767]

[-0.35196482]

[-0.24494025]

[0.15505975]]

1.5

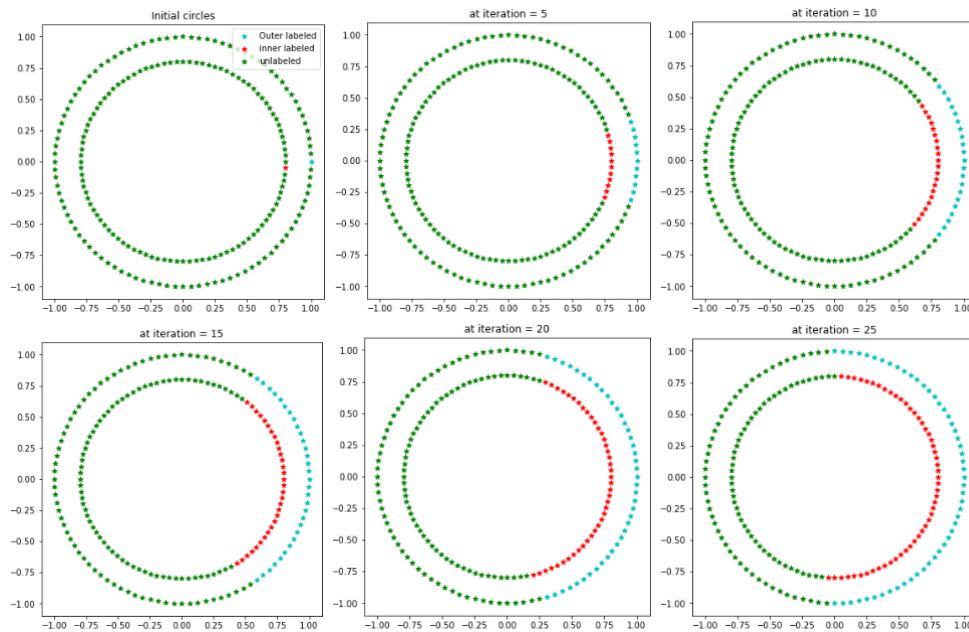
$$F_u = -L_{uu}^{-1} L_{uu} Y_1$$

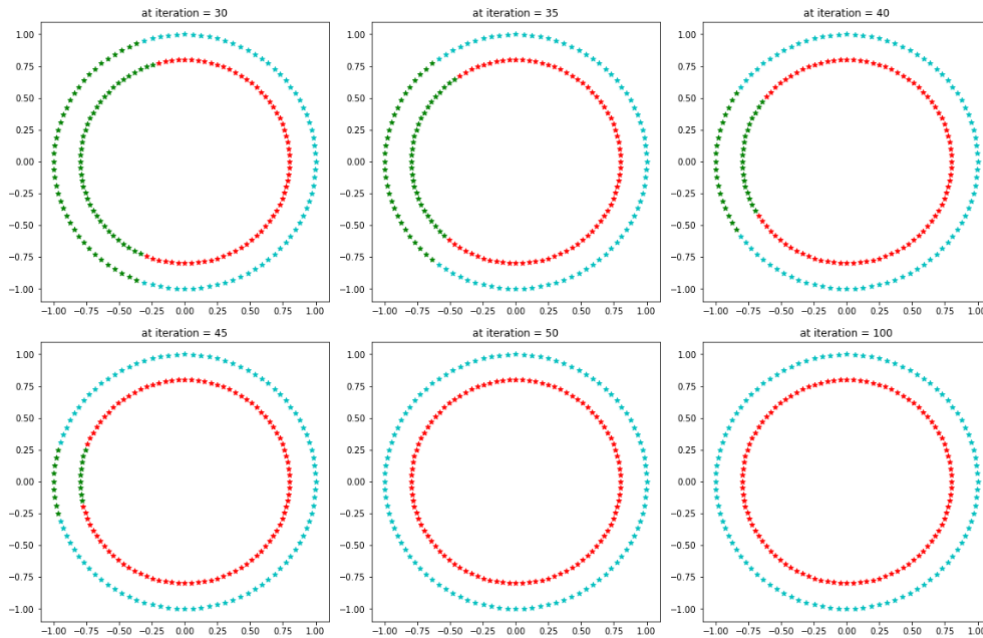
$$F_u = \begin{bmatrix} 0.38461538 & 0.61538462 \\ 0.15384615 & 0.84615385 \\ 0.07692308 & 0.92307692 \end{bmatrix}$$

$$Y_u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

As you can see, every data point has been assigned a second label. Furthermore, we see that, in contrast to the label spreading method, which modifies the label nodes, the algorithm can identify the unlabeled points without affecting the labeled nodes.

Problem 2

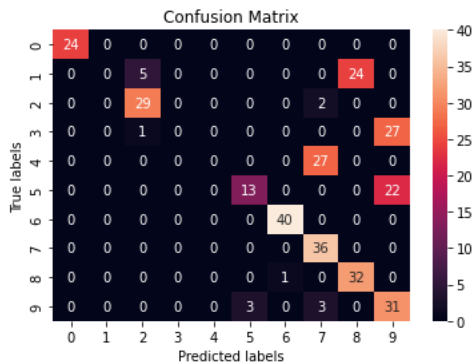




From the above results we can see that the labels are spread correctly, and model has converged at 50 iterations.

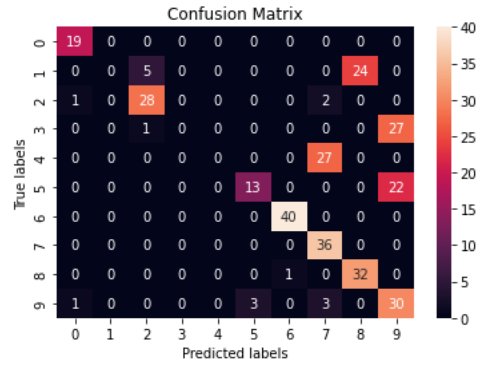
Problem 3

Label propagation model: 10 labeled & 320 unlabeled points (330 total)



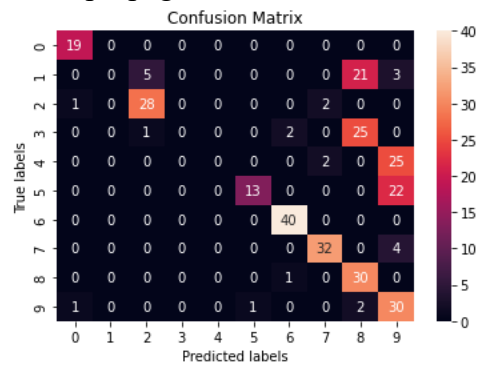
Accuracy: 64.1%

Label propagation model: 15 labeled & 315 unlabeled points (330 total)



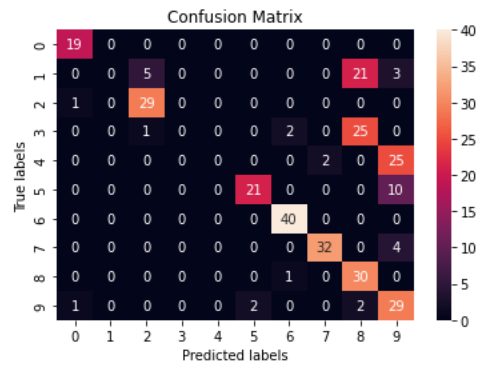
Accuracy: 62.9%

Label propagation model: 20 labeled & 310 unlabeled points (330 total)



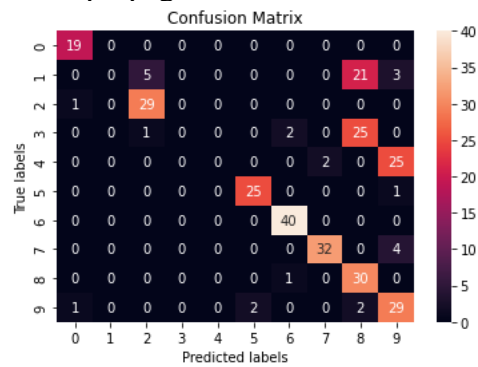
Accuracy: 61.9%

Label propagation model: 25 labeled & 305 unlabeled points (330 total)



Accuracy: 65.6%

Label propagation model: 30 labeled & 300 unlabeled points (330 total)



Accuracy: 68.0%

Problem 4

The test results for the labeled data with the numbers [60, 120, 180, 240, 300] are as follows:

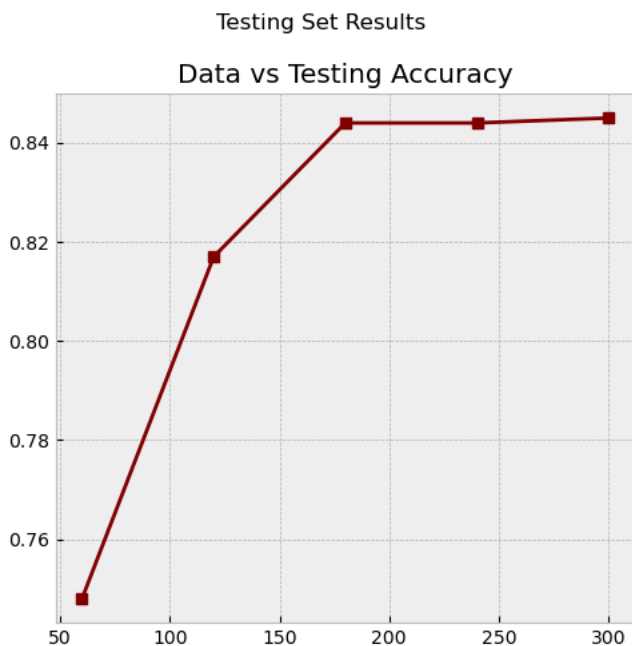
Accuracy= 0.7480

Accuracy= 0.8170

Accuracy= 0.8440

Accuracy= 0.8440

Accuracy= 0.8450



The testing accuracy rises with the amount of labeled data and peaks at about 84.5% as we increase the label data.