BigML-Assignment 6
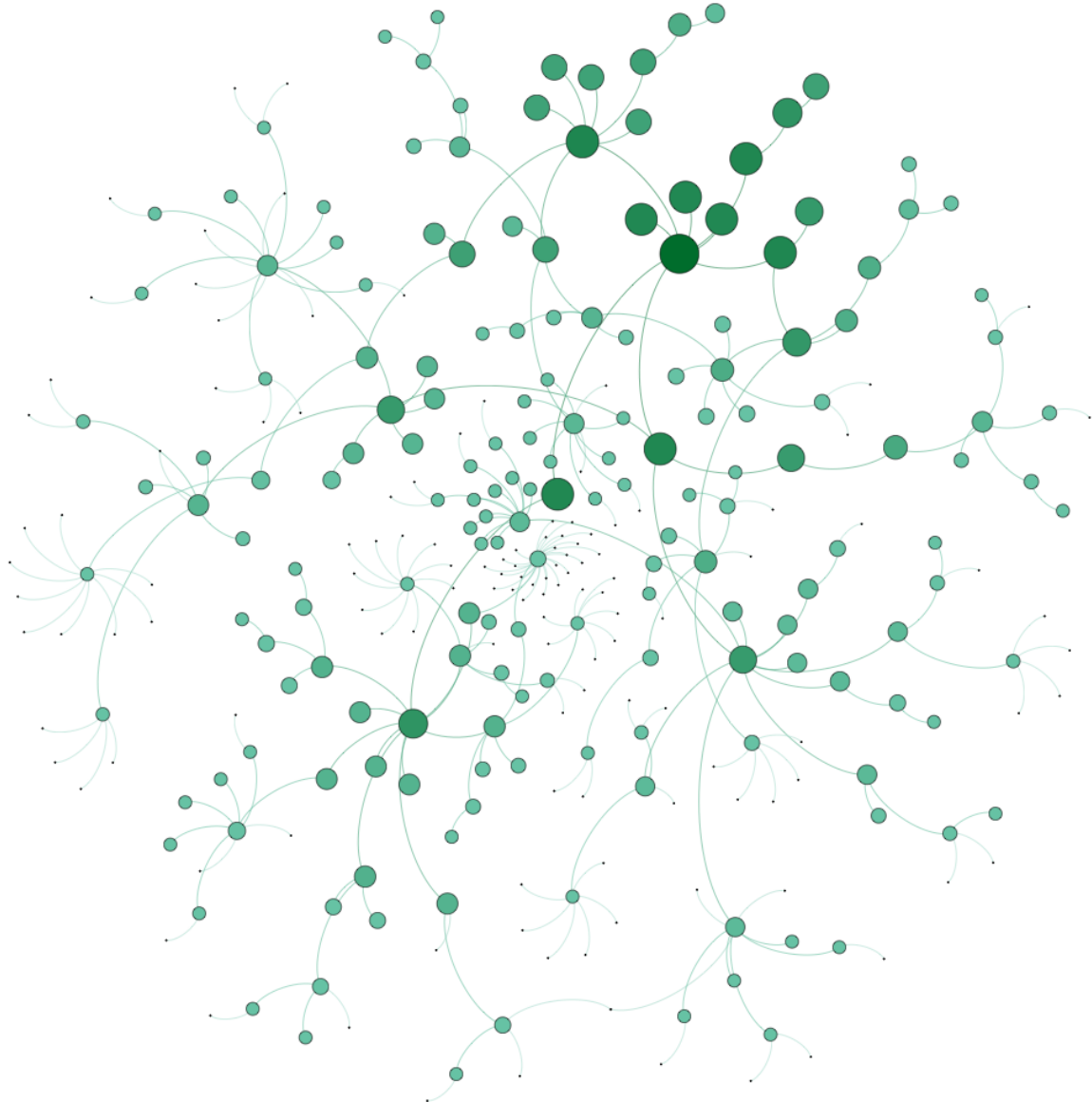
Chen Sun / chens1

1. Help Policy

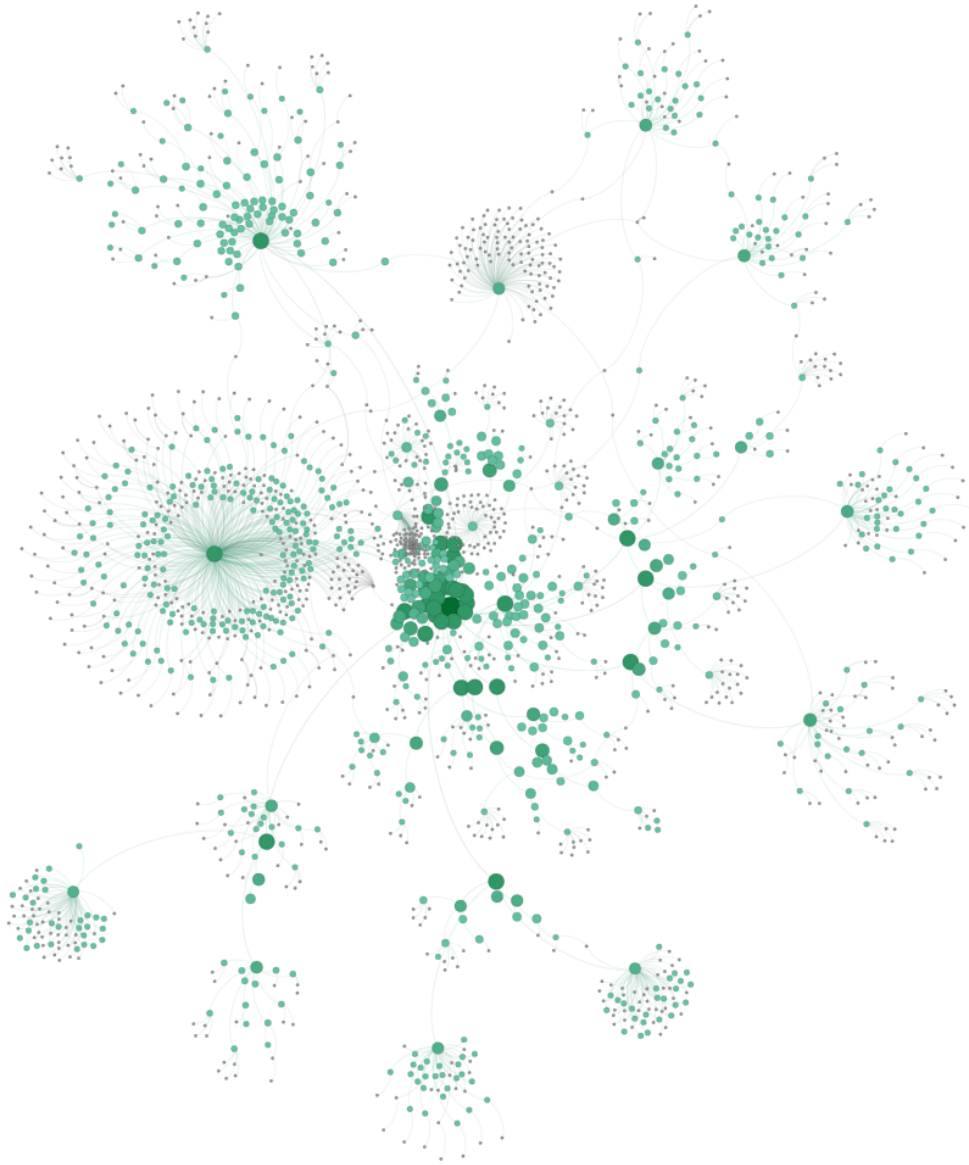No.

No.
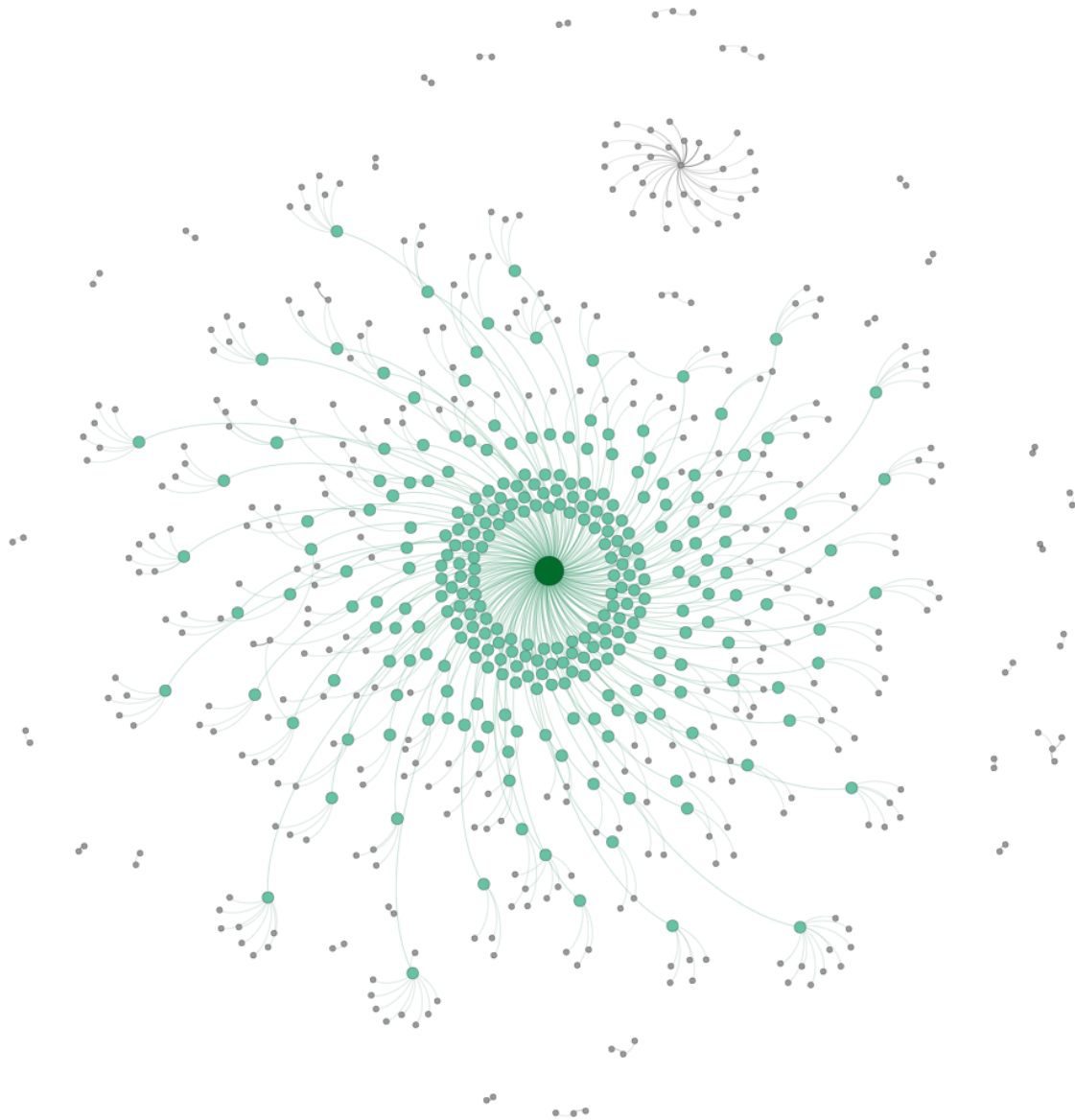
2.

This is the visualization of Machine_learning. With the center being Machine Learning.

3.

This is the visualization of **Carnegie_Mellon_University**, center being the seed and the circle being the Pennsylvania. Alpha = 0.3 and epsilon = 0.00001

This is a visualization of **Pennsylvania,** alpha = 0.3 and epsilon = 0.00001, it is very beautiful and symmetric.
3.
Here is a series of experiment I did on the data.

alpha	epsilon	Total Node	Runtime (s)
0.3	0.000001	426	15.59
0.3	0.00001	177	12.405

0.3	0.0001	71	10.001
0.3	0.001	25	6.898
0.3	0.01	8	4.286
0.3	0.1	1	1.717

alpha	epsilon	Total Node	Runtime (s)
0.15	0.00001	427	24.05
0.2	0.00001	293	21.011
0.25	0.00001	212	16.059
0.3	0.00001	177	12.405
0.35	0.00001	145	11.599
0.4	0.00001	124	9.779
0.45	0.00001	98	8.349

The first series change the value of epsilon, changing in 10 times each. Although the epsilon change exponentially, the runtime changes linearly, indicating linear decrease in scanning time. And the final total node decrease exponentially but not as fast as the epsilon value.
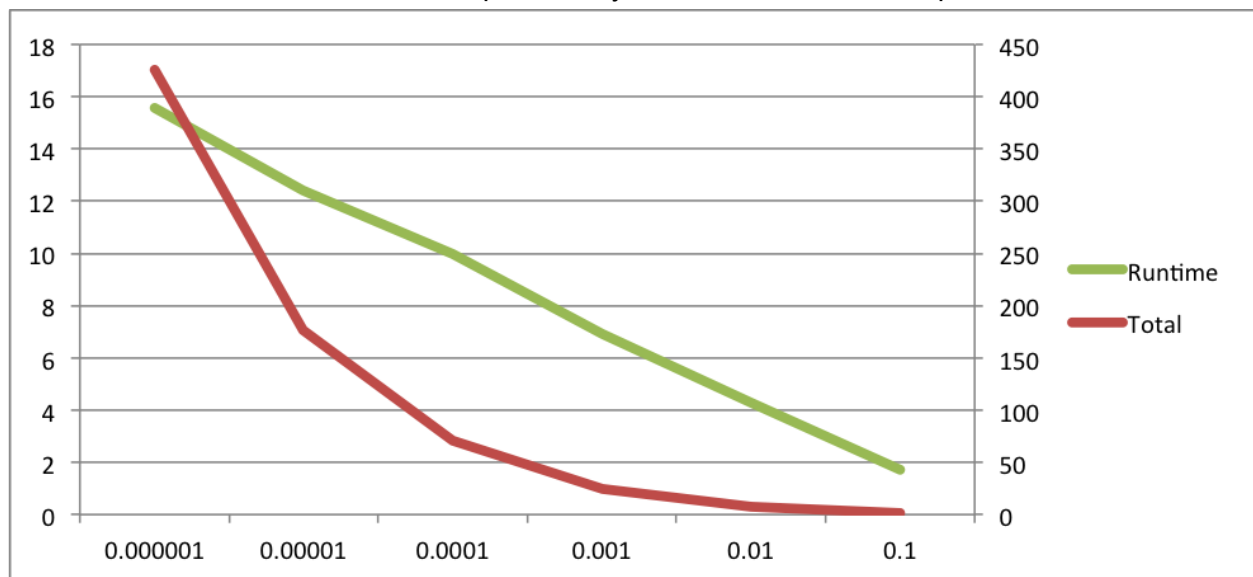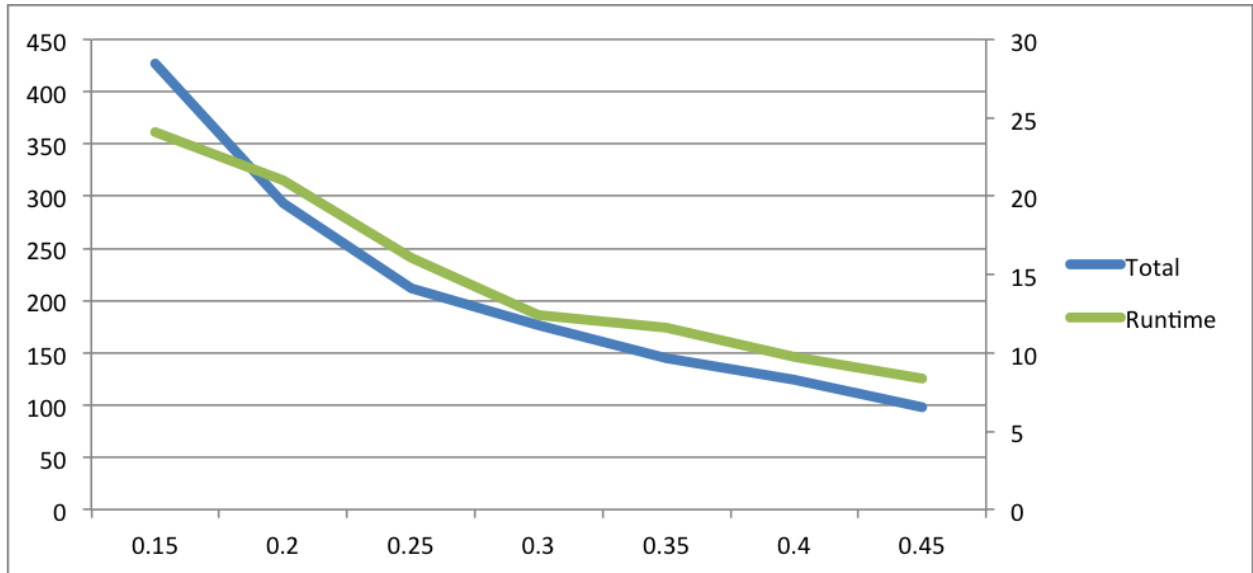
The alpha value changes linearly, and correspondingly, the total nodes retrieved and runtime decreases linearly.

That means the result graph is more sensitive than the epsilon value. Since the alpha value is usually assigned with larger weight in calculating scores, and epsilon only affects those at the very small value stage. Thus alpha is more important in affecting the score. As is shown in the test.