Example 1:

 

The results for the L2 version and L1 version looks like below. (I used 20 nearest neighbor to construct the graph). While for both problems, the classification is perfect, l1 version leads to something that succeeds with much less confusion.



**Example 2**



The classification is great from just a few random labels for both l1 and l2. But there is a difference between the two in the solution. The l1 version seems to end up with a pretty nice piecewise constant signal.



We will see if the L1 version would work better in real datasets,.