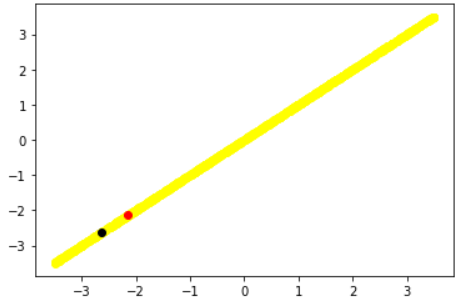
1. Simple L1 regression with only one parameter .
2. Line

The simulated line includes 1000 data points. Some points are uniformly selected from dataset and the points inside its neighborhood with given radius are chosen to do L1 regression. Then we observe the points with nonzero coefficients. The radius is controlled by parameter . Under each combination of parameters, 200 points are uniformly selected which is 20% of the population.

The result is very sparse, less than 3 points in each neighborhood are connected with central point. And the distances between central point and connected points are quite large. This could be due to the relatively large radius of the neighborhood. Like this



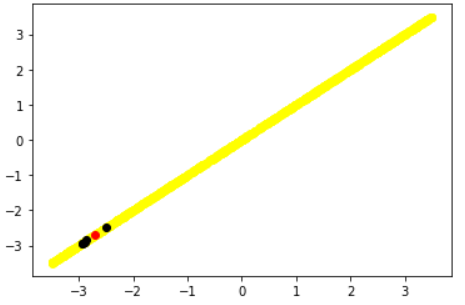
Under this situation, the results are also sparse as previous part. But due to the smaller radius, the distance between connected points are much smaller.



Under this situation, the result is similar to the previous part.

In these cases, the results are similar. The connected points are close to central point but only few points are connected, less than 7.

In this case, more points are taken into regression because of the increasing radius and the connections are little more than the previous situation. But still too sparse like this.

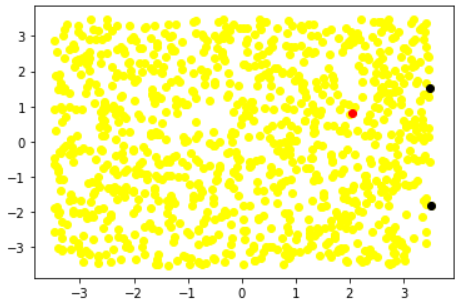
Points are closer but very few are connected.

To sum up, in a line, larger radius and smaller lead to more connections. Larger radius and larger lead to larger distance between connected points. Thus if we want to preserve the properties of KNN in a single line, we need proper radius and smaller .

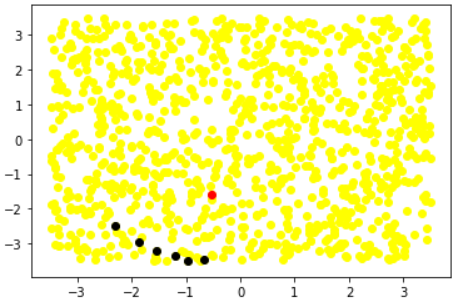
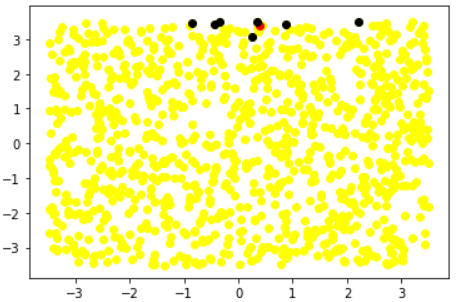
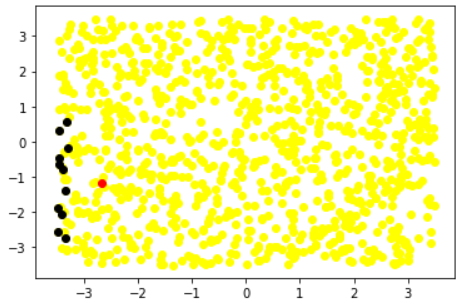
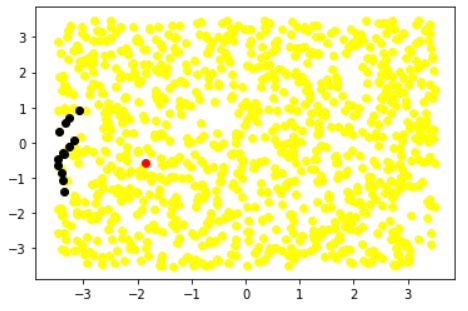
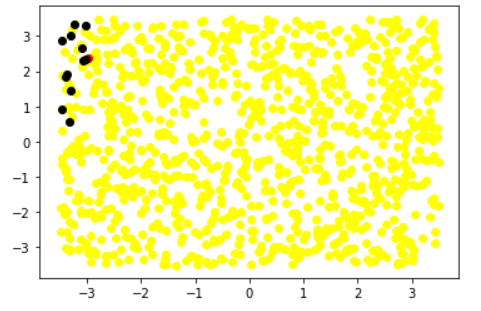
1. Plane

The simulated plane includes 1000 data points. 200 points are uniformly selected from dataset and for each point, L1 regression will be operated on points inside its neighborhood. And we will check the distribution of these points and their coefficients.

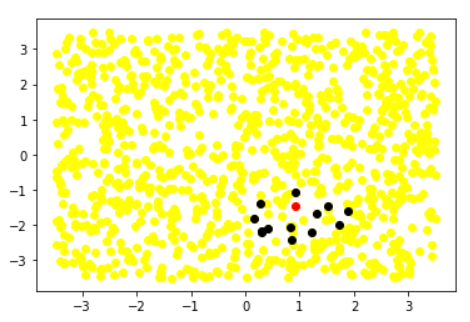
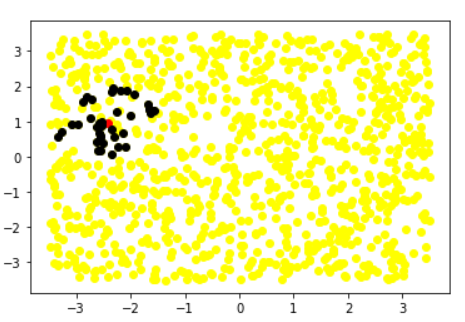
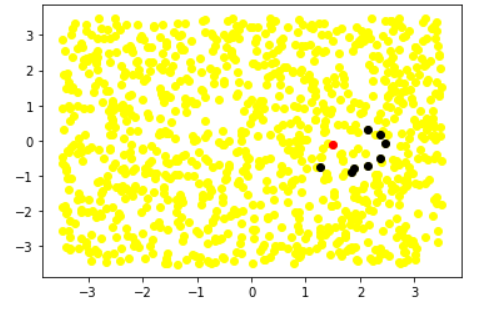
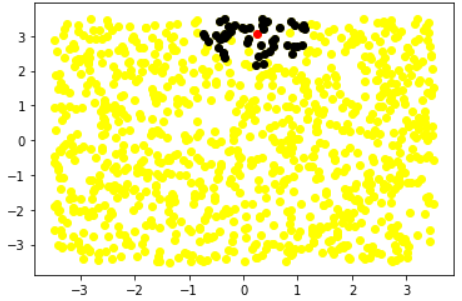
This is really large that the whole dataset will be involved into regression. Since is large too, less than 3 points will be connected with central point which is a bad result.

As we could see, points are far away from each other and we could not see any linearity. This indicates the importance of local L1.

We decrease the radius and , more points are connected, up to 13. And the distribution of points are different, like below.

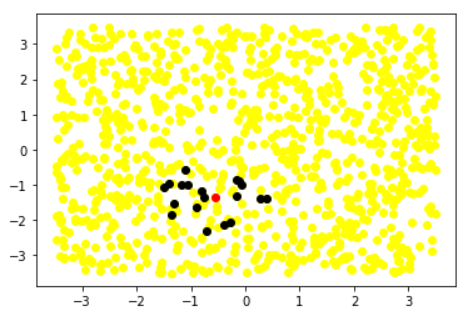
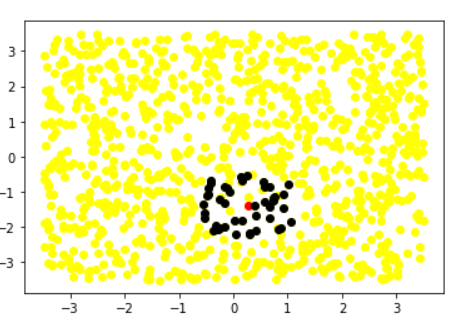
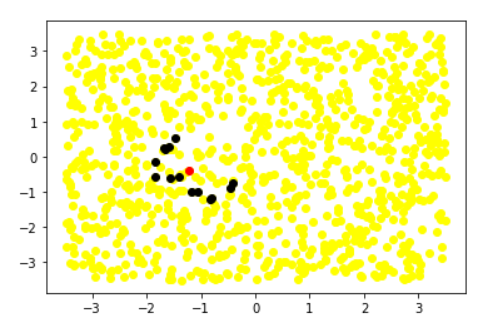
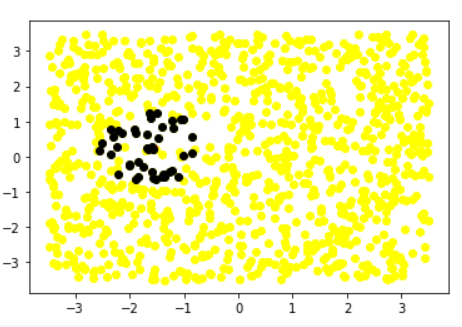
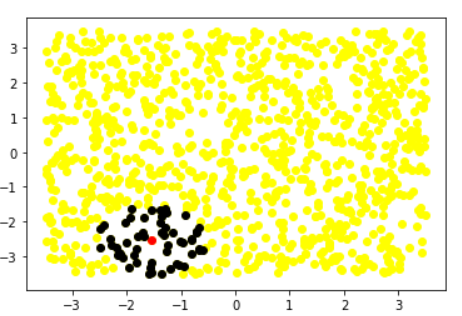
From these plots, we could see linearity among connected points but we could not discover any sign of neighborhood among these points which totally oppose the KNN. And an interesting fact is these points tend to be close to boundaries. Like shown in 5 plots above, black points are all quite close to the boundary which is nearest to the red point. These points are sparse and not close to central point. This might be because the L1 penalty dominates the optimization and the neighborhood is too big.

In this case, some dramatic changes happen. We start to discover the neighborhood from results. And unlike any previous results, much more points are connected. Most points are connected to half of their neighbors and the largest amount is 72 which is absolutely not sparse and shows more similarity with KNN.

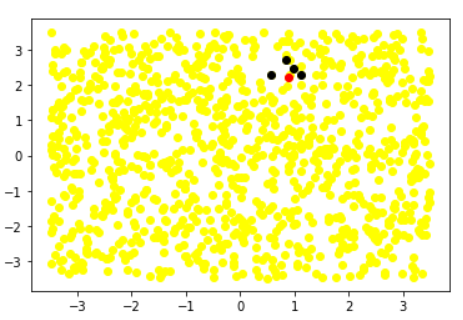
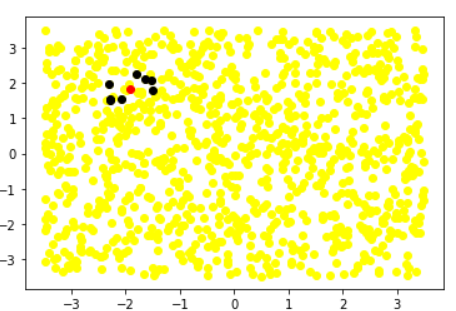
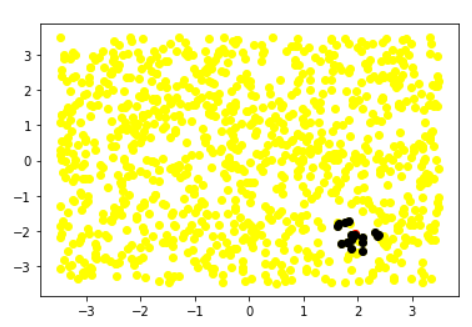
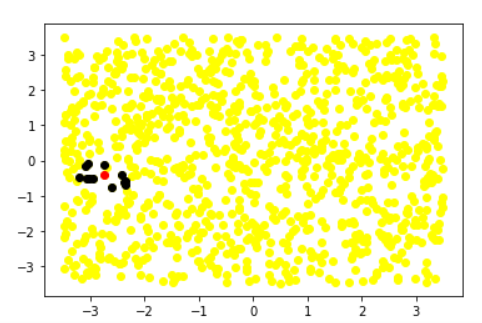


The second plot looks a little bit different because the fewer points and larger distance. That happens on some samples, but I am not sure why this happens, maybe this is because the distribution and density of points in this neighborhood.

We keep decreasing , and similar to previous part, an obvious neighborhood could be observed.

 The last plot is a kind of exception but there are very few exceptions. The decrease of do improve the performance of local L1.

Under this circumstance, the amount of connections is decreasing because of the shrink of neighborhood. Most points are connected to 7-15 neighbors which is much less than the previous situation.



From plots, obvious neighborhood could be observed but fewer neighbors. And due to the small radius, connected points are close to central points which is nice.

To sum up, when is small such as 0.0001, we will discover obvious neighborhood. And when setting radius to be proper, more than half neighbors will be connected which preserve the property of KNN really well(since we are testing on simple plane).

1. Cube

1000 points are simulated and 200 samples are selected uniformly. For each sample, its neighbors in the neighborhood are used for a regression and we check the connections and coefficients. Since this is a 3 dimension object, it is unclear to see points in a 3d view thus we prefer to observe the object in 3 dimensions separately.

We choose large and large at the first time. This is like doing L1 globally and we will see what happened. As we thought, very few points are connected, less than 4. We do not have to discuss more because this is not what we want to see.

