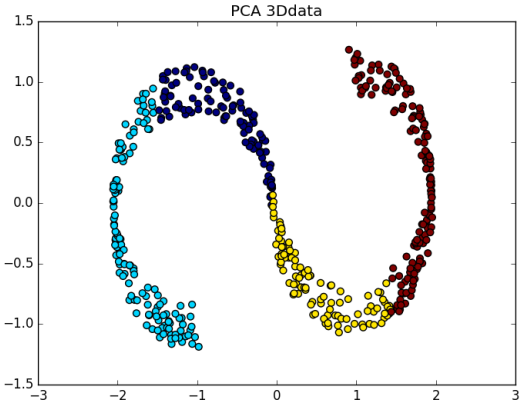
Machine Learning

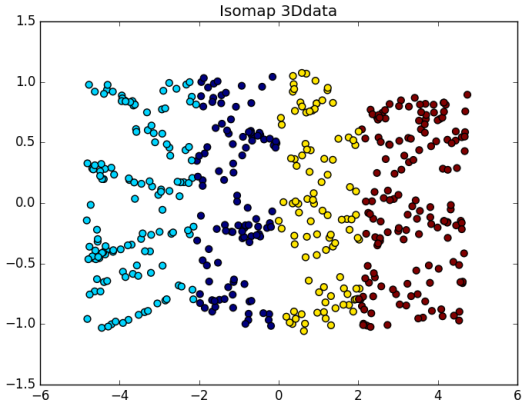
Roberto Fernandez

Problem Set 3

Exercise 1

Below we compare our plots for PCA and Isomap dimensionality reduction on the 3 dimensional swiss roll.





We can notice very stark differences between the plots for PCA and Isomap. The greatest difference arises due to the linear nature of PCA while Isomap has no such restriction. Since the data we are trying to run dimensionality reduction on is a swiss roll in $3-$D then the non-linear nature of the structure remains while Isomap is able to remove this and we end up with a much more uniformly distributed collection of points while PCA still has a very apparent "wavy" shape.

Exercise 2

We have the below average error plot:

It is thus clear that the error is monotonically decreasing after each run through the data, which makes sense, and we stop at $M = 5$ since it gives a reasonable prediction error rate of around 5%.

Exercise 3

We have the below average error plot:

It is thus clear that the error is monotonically decreasing after each run through the data, which makes sense, and we stop at $M = 5$ since it gives a reasonable prediction error rate of around 5%, similar to the above case.