

Theory Exercise 3

1) What are the advantages of t-SNE over PCA?

Advantages with t-SNE

1. It handles non linear data good. This is because t-SNE places similar data points together and dissimilar are kept far apart.
2. It also preserves local and global structure. AKA points clustered together in high-dimensional structure will be close together in a lower structure. It also do the opposite for points far apart in the high-dimensional structure.

Disadvantages with t-SNE compered to PCA

1. Uses a lot of computer power because it is $O(n^2)$ (I think so at least). If you have a lot of features use PCA to get it below 10'000 and then use t-SNE.
2. It varies every run.
3. Has hyperparameter that might need tuning.
4. Can find patterns in the random noise. So needs to run a few times with different sets of hyperparameter.

Consider three points a,b and c.

I will list them in the order of lowest loss to the highest.

1. a and b are close to each other, and c is far away from them.
2. a, b and c are all close to each other.
3. a, b and c are all far away from each other.
4. a is far away from both b and c, that are close to each other.

1

This will give the lowest possible loss because they are keeping the "original" distances to each other.

2

Being close to something you should not be dose not give that big of a loss compered to being far away of something you should not.

3

a and b are faraway from each other and that gives a bigger loss then 2.

4

a and b are still faraway from each other but c is close to b and that gives additional loss. So 4 has the biggest loss of all of these configurations.

Link to paper that follows the same structure as the video

<https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>