1. Theoretical Tasks

In this exercise we will focus on visualizing and understanding what the Con- volutional Neural Networks we have discussed in the previous weeks are doing under the hood. Watch the following presentation on t-SNE, and answer the questions below:

- https://www.youtube.com/watch?v=RJVL80Gg3lA

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

ANSWER: t-SNE has a bunch of advantages over PCA, for example t-SNE is non-linear which means it captures non-linear relationships between variables in the data. That lead to more effective vizualisation of complex patterns in high-dimensional data.
Also t-SNE preserves the local structure of the data which lead to preserve better the separation between these groups, as we see with the second model.
t-SNE is less sensitive to outliers in the data than PCA. It's due to the uses of a probability distribution to measure the similarity between data points, which downweights the influence of outliers.
The output of PCA is a set of orthogonal axes that represent the directions of maximum variance in the data. While this is useful for reducing the dimensionality of the data, it can be difficult to interpret these axes in terms of the original variables. In contrast, the output of t-SNE is a two-dimensional (or three-dimensional) map that can be easily visualized and interpreted.

2) Consider three points a, b and c in the high-dimensional space.

Let us assume that a and b are very close to each other, and that c is very far away from both of them. How would the loss behave if the following is true for the low-dimensional space? Feel free to use dummy numbers to support your answer.

- a, b and c are all close to each other.

ANSWER: loss low

- a and b are close to each other, and c is far away from them.

ANSWER: loss middle

- a, b and c are all far away from each other.

ANSWER: loss high

- a is far away from both b and c, that are close to each other.

ANSWER: loss high