CSCC11 ASN1

▼ part 1

- Mean Absolute Error (MAE) & Mean Squared Error (MSE). Define these metrics and discuss why one would be preferred over the other.
    - Mean squared error is more preferred because it is more sensitive to the outliners. It can better detect outliners. We squared the error, which means if a data point has larger error (i.e. is an outliner), it will have more weight on the final result (i.e. the mean error) for the whole data set.
- Complete the code to determine the Silhouette Coefficient associated with each value of K ranging between 2 and 10. Which value would be the most appropriate?
    - 2 will be the most appropriate
- Provide a brief discussion regarding the factors that might have contributed to this result.
    - different input and arguments for `KMeans` will result in a different error outcome. For example, if there are more outliners in our input data, the error will get larger.

▼ part2

- Examine what happens when you change the distance between RBF centres. What is the effect of RBF width on the reconstruction?
    - The image is not fixed well when we set `spacing=1`, but other values from 2 to 9 seems not much difference. The change of width will change the output image. Therefore, we need to test to find out the optimal value of width.
- What can you say about tuning the model as you increase the amount of hyper-parameters?
    - By increasing the amount of hyper-parameters, it is possible to build a better model. For example add a smooth term λ, we can avoid overfitting. However, we need to set the appropriate value of our hyper-parameters.