CSCS933/433: Assignment 2:
A study of non-linear dimensioanlity reduction methods

Marking scheme (Rubric)
Maximum of 165 marks accounting for 25% of total assessment marks

May 2, 2020

Abstract

The abstract should be informative and provide a summary of key results and insights.

1. Maximum of 5 marks

2. Minimum of 1 mark

1   Introduction

The introduction should provide the reader with an informative background to dimensionality reduction; what it is; why it is important; overview of what it entails.

1. Maximum of 10 marks

2. Minimum of 2 marks

2   Theory and properties of techniques

This section should provide the general theoretical underpinning of dimensionality reduction. You required to use your knowledge of linear algebra to explain the theory of dimensioanlity reduction

1. Maximum of 5 marks

2. minimum of 1 mark

The subsections that describe each method is worth 40 marks (10 marks each). Your description should exhibit your understanding of the theory using appropriate mathematical formulations.

2.1   Principal Component Analysis (PCA)

2.2   Kernel PCA

2.3   Autoencoders

2.4   Local linear embedding

3   Experiments

1. Maximum of 5 marks

2. Minimum of 1 mark

In this section you should clearly describe the experiments you have carried out and what you expect to learn from the results.

3.1   Experimental setup

This section should describe the how you set up the experiments including the parameters you used in each experiment. The rationale for the choice of each parameter should be provided.

1. Maximum of 10 marks

2. Minimum of 1 mark

3.2   Artificial datasets

1. Maximum of 15 marks

2. Minimum of 2 mark

This section should describe the manner of generating the artificial datasets to be used in the experiments. Images showing the actual dataset generated from your code should be inserted and commented upon. For example you may want to explore how the parameters affect the dataset generated. if you have performed any pre-processing on the data, it should mentioned and described here along with the rationale.

3.3   Natural datasets

1. Maximum of 15 marks

2. Minimum of 2 mark

This section should describe the characteristics or each natural dataset used in the experiments. Images showing samples of the actual dataset should be inserted and commented upon. If you have performed any pre-processing on the data, it should be mentioned and described here along with the rationale.

3.4   Results - Artificial datasets

1. Maximum of 15 marks

2. Minimum of 2 mark

This section should clearly describe your results on the artificial dataset along with appropriately formatted tables and any other diagrams you deem appropriate.

3.5   Results - Natural datasets

1. Maximum of 15 marks

2. Minimum of 2 mark

This section should clearly describe your results on the natural dataset along with appropriately formatted tables and any other diagrams you deem appropriate.

4   Discussion

1. Maximum of 15 marks

2. Minimum of 2 mark

This section should provide an in-depth discussion of the results and their implications for classification tasks. You should also discuss the implications of your results in terms of generalizations. You are required to think about the possibility of generalizing the results of your experiments. What is required (if any) to be able to generalize your results beyond the case of the datasets you have used?

5   Conclusion

1. Maximum of 10 marks

2. Minimum of 1 mark

What reasonable conclusions can you reach, based on your results?

The reference section will attract a maximum of 5 marks. This will be based on correct formatting of the listings and the number of extra and appropriate references cited.

References

Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel methods for pattern analysis*. Cambridge: Cambridge Univerrsity Press.

van der Maaten, L. J. P., Postma, E. O., & van den Herik, H. J. (2008). *Dimensionality reduction : A comparative review*. online. Retrieved March 2020, from `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.5472&rep=rep1&type=pdf`

van der Maaten, L. J. P., Postma, E. O., & van den Herik, H. J. (2009). *Dimensionality reduction : A comparative review*. online. Retrieved March 2020, from `https://lvdmaaten.github.io/publications/papers/TR_Dimensionality_Reduction_Review_2009.pdf`