Tests & Quizzes

Worksheet 06 - Dimension reduction challenge

Return to Assessment List

Part 1 of 1 - 10.0 Points

You are given a **three-dimensional time series** data: $\mathrm{data\_x} \in \mathbb{R}^{T\times 3}$, with $T = 100,000$ (number of data points). The data points in this time series can be grouped/clustered into four (metastable) states which, unfortunately, cannot be separated by simple geometric means.

Your first task is to perform a dimension reduction from the original three-dimensional representation into a **one-dimensional** time series in such a way that **the four different states become disentangled**. In other words, the four states must be easily separable in the one-dimensional representation. We recommend to use a **time-lagged autoencoder** for this task, but your are free to try other approaches, too.

Hint: a proper data whitening might be very beneficial for the training.

In the second task, you have to discretize your one-dimensional representation, i.e., all data points which seem to belong to the same (metastable) state must be assigned the same label $\mathrm{data\_y} \in \{0, 1, 2, 3\}$. The order in which you label the metastable states is not important (we will test all combinations).

Hint: you can do the discretization, e.g., by defining four different positions (cluster centers) within your one-dimensional dataset and assign each data point to the closest cluster center.

For validation purposes, we have added a validation set containing 1000 data points ($\mathrm{validation\_x}$) and their corresponding metastable state labels ($\mathrm{validation\_y}$).

Attachments

dimredux-challenge-01-data.npz 2255 KB

Question 1 of 1

10.0 Points

Store the label prediction ($\mathrm{data\_y}$) in a one-dimensional numpy file (**prediction.npy**) with an *integer data type (e.g. np.int32)* and upload your file. The shape should be (100000,).

Please note that you need to use exactly the shown file format, file name, and array shape as described above. Otherwise, we might not be able to correctly process your submission.

Please note that all members need to submit their own prediction or they won't be awarded points.

Please also note that this question is manually graded and that you should ignore any automatically set result until a teaching assistant had the opportunity to look at your submission.

The points will be awarded according to the label accuracy in your submission according to the following scheme:

accuracy	points
$\geq 95\%$	10
$\geq 90\%$	7
$\geq 80\%$	5

Please remember actually submitting your solution (and make sure it is the right file)!

prediction.npy (400.13 KB)

- Site Browser
- Lehr- und Studiumsplanung
- Vorlesungsverzeichnis
- Accessibility Information

- Powered by Sakai
-