

MLCMS Final project - One Page summary

(Names of the group members)

Learn and visualize representations for large data sets

Project Summary

Our project involves understanding Manifold Learning(ML) algorithms by using the the new `megaman` Python package used for scalable manifold learning. We use different classes of **embeddings** to study a system generated dataset and also a self generated dataset.

Task 1/5 - Description of Dataset used in the paper

Analyse the design philosophy behind `megaman`, describe the different embeddings available to use. Download and Install `megaman`. Describe `make_swiss_roll()` function and dataset generted. **(Note that this is not possible anymore)**

Task 2/5 - Implementation of the fast Representation algorithms

Outline tasks performed by generic manifold learning algorithm. Implementing the different embedding algorithms, like `SpectralEmbedding`.

Task 3/5 - Testing implementation on large self-created dataset

Using `SpectralEmbedding` to analyse and test self created dataset.

Task 4/5 - Testing implementation on large dataset from the paper

Using `SpectralEmbedding` to analyse dataset created by `make_swiss_roll()` function.

Task 5/5 - Comparison of results between the 2 datasets

Analyse the variation in dataset. Explain how the algorithm processed the two datasets differently. Compare visual results from both datasets, for comparison.
