Manifold Learning and Data Visualization

Heidelberg, September 25th, 2018

data2day 2018

Dr. Stefan Kühn

https://www.xing.com/profile/Stefan_Kuehn46

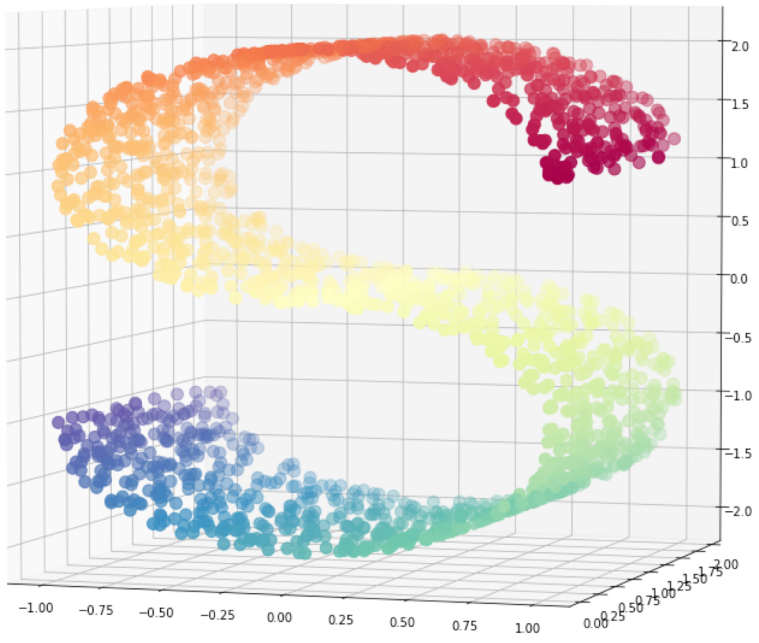
https://www.linkedin.com/in/stefan-k%C3%BChn-020a34119/

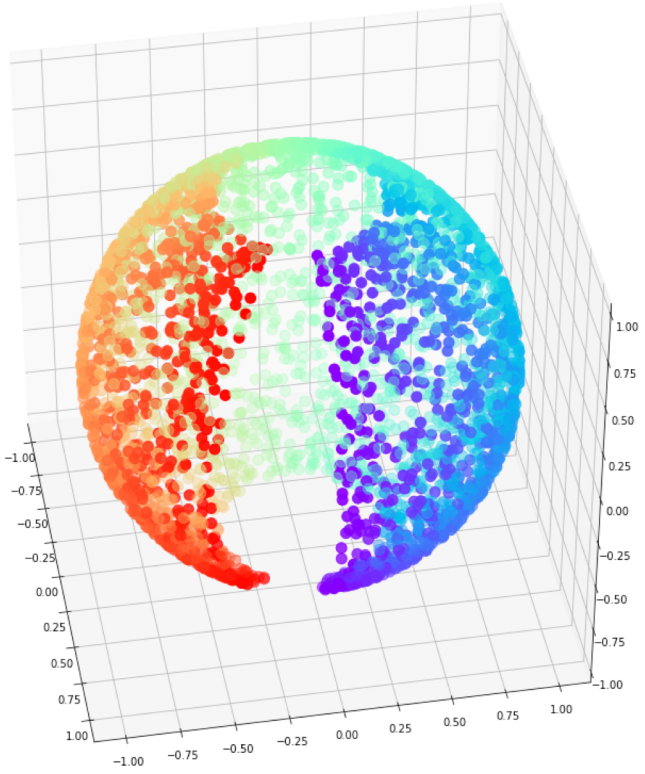
https://de.slideshare.net/StefanKhn4

XING

What is a manifold?

Mathematical concept from Differential Geometry





What are properties of a manifold?

Important Properties – Topology and more

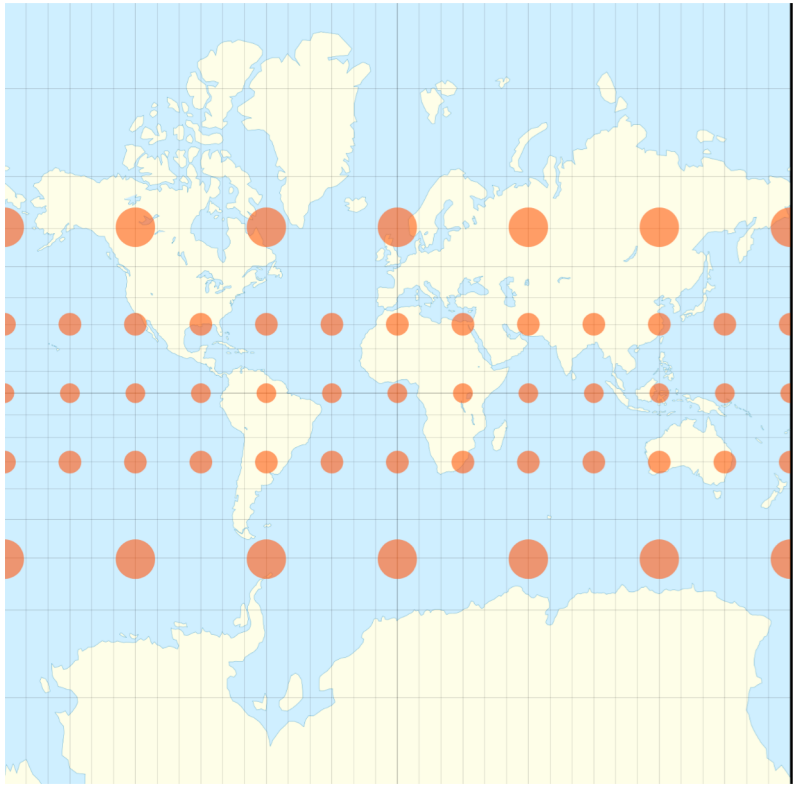
- Number of Connected Components
- Holes
- Curvature
- Smoothness
- Dimensionality
- ...you_name_it...

What are properties of a good visualization?

Preserve important properties

- Number of connected components?
- Holes?
- Curvature?
- Smoothness?
- Dimensionality?
- Distances between points?
- Angles, orientations?
- Local versus global properties?

You cannot have it all!



Manifold Learning Methods in sklearn

- **Locally Linear Embedding**
  - Neighborhood-preserving
- **Isomap**
  - Quasi-isometric
- **Multi-Dimensional Scaling (MDS)**
  - Quasi-isometric
- **Spectral Embedding**
  - Spectral clustering based on similarity
- **T-Distributed Stochastic Neighbor Embedding (tSNE)**
  - Preserves probabilities
- **Local Tangent Space Alignment (LTSA)**

Demo Time

Sometimes, words are insufficient...

In mathematics, a **manifold** is a topological space that locally resembles Euclidean space near each point. More precisely,

One-dimensional manifolds include lines and circles, but not figure eights (because they have *crossing points* that are not l
self-intersections) in three dimensional real space, but also the Klein bottle and real projective plane, which will always self-

But then God gave us code!

```
print(__doc__)

from time import time

import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
from matplotlib.ticker import NullFormatter
%matplotlib inline
from sklearn import manifold, datasets
```

Resources

Scikit-learn documentation

http://scikit-learn.org/stable/modules/manifold.html

http://scikit-learn.org/stable/auto_examples/manifold/plot_compare_methods.html

http://scikit-learn.org/stable/auto_examples/manifold/plot_manifold_sphere.html

http://scikit-learn.org/stable/modules/random_projection.html

http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

Github repo with worked examples

https://github.com/cc-skuehn/Manifold_Learning

Jupyter Lab

https://jupyterlab.readthedocs.io/en/stable/index.html

Citation

Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html

Thank you!

Heidelberg, September 25th, 2018