

PR1/2 Demo

Christian Permann

Faculty of Computer Science, University of Vienna,
Währinger Straße 29, 1090 Vienna

13.11.2018

Milestones P1 updated

- ▶ Create a basic visualization tool. (done)
- ▶ Define interface for data manipulation. (done)
- ▶ (Be compatible with ELKI.)
- ▶ Allow data import. (CSV/Arff available, done)
- ▶ Create data generation logic. (ELKI generator available, possibly add simple GUI with reduced functionality)
- ▶ Create dim-reduction for visualization and implement it. (PCA and T-SNE, done)
- ▶ Allow to generate a scatter plot matrix. (done)

The Tool

Demo

The Tool - Dimensionality Reduction Interface

```
public interface IDimensionalityReduction {  
    JPanel getOptionsPanel();  
  
    String getName();  
  
    boolean reduce(PointContainer container);  
}
```

Figure 1: The interface for dimensionality reduction

Papers for PR2

- ▶ LineUp: Visual Analysis of Multi-Attribute Rankings [1]
- ▶ WeightLifter: Visual Weight Space Exploration for Multi-Criteria Decision Making [2]
- ▶ Metric Factorization for Exploratory Analysis of Complex Data [3]
- ▶ DimStiller: Workflows for dimensional analysis and reduction [4]
- ▶ Comparing clusterings: an axiomatic view [5]
- ▶ Comparing subspace clusterings [6]
- ▶ External evaluation measures for subspace clustering [7]

Ideas for evaluating Clusterings

- ▶ different quality measures as described in [7]
- ▶ a weighted average of measures like in ClusterVision[8]
- ▶ using [1, 2, 3] as basis for analyzing the quality of measures and deciding on different weights (kind of a better reasoning for the choice compared to clustervision)
- ▶ **clustering (OPTICS?) clusterings and visualizing groups of results across multiple algorithms and settings (new idea?) using [5, 6]** in regards to the distance measure
- ▶ possibly training a weighted average of measures via supervision with a Neural Network (needs lots of data and known optimal clusterings; generalizeable result?)

My Idea

Using OPTICS to cluster clusterings could result in a reachability plot that hierarchically shows groups of clustering results that agree on the result. The output of the algorithm can also be visualized in a symmetric heat-map as shown in our OPTICSVis Project. Here it may be possible to see which clusterings overlap with others and which are vastly different, in a hierarchical view. One problem here may be though, that the distance measure should satisfy the triangle inequality for useful measures(, I think).



S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit, “Lineup: Visual analysis of multi-attribute rankings,” *IEEE Transactions on Visualization and Computer Graphics (InfoVis ’13)*, vol. 19, no. 12, pp. 2277–2286, 2013.



S. Pajer, M. Streit, T. Torsney-Weir, F. Spechtenhauser, T. Möller, and H. Piringer, “Weightlifter: Visual weight space exploration for multi-criteria decision making,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, pp. 611–620, January 2017.



C. Plant, “Metric factorization for exploratory analysis of complex data,” in *2014 IEEE International Conference on Data Mining*, pp. 510–519, Dec 2014.



S. Ingram, T. Munzner, V. Irvine, M. Tory, S. Bergner, and T. Möller, “Dimstiller: Workflows for dimensional analysis and reduction,” in *IEEE Conference on Visual Analytics in Science and Technology (VAST)*, pp. 3–10, October 2010.

(26 out of 94 accepted).



M. Meilă, “Comparing clusterings: an axiomatic view,” in *In ICML '05: Proceedings of the 22nd international conference on Machine learning*, pp. 577–584, ACM Press, 2005.



A. Patrikainen and M. Meila, “Comparing subspace clusterings,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, pp. 902–916, July 2006.



S. Günnemann, I. Färber, E. Müller, I. Assent, and T. Seidl, “External evaluation measures for subspace clustering,” in *Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM '11*, (New York, NY, USA), pp. 1363–1372, ACM, 2011.



B. C. Kwon, B. Eysenbach, J. Verma, K. Ng, C. D. Filippi, W. F. Stewart, and A. Perer, “Clustervision: Visual supervision of unsupervised clustering,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, pp. 142–151, Jan 2018.