**Applying Dimension Reduction Techniques to Measured Cardiac Blood Flow Data**

**Team Members:**

1. Jayanth Varma (student id: 225821, study course: M-DKE, current semester: 4, email: [jayanth.dantuluri@st.ovgu.de](mailto:jayanth.dantuluri@st.ovgu.de)
2. Manish Sreerangam (student id: 225856, study course: M-DE, current semester: 4, email: manish.sreerangam[@st.ovgu.de](mailto:jayanth.dantuluri@st.ovgu.de))

**Background:** The aorta is the largest vessel in the human body and often affected by different diseases. A typical disease relates to the aortic valve, called bicuspid aortic valve. Researchers found out that there appear to be relations between hemodynamics and cardiac pathologies. Patient-specific blood flow information can be measured by four-dimensional phase-contrast magnetic resonance imaging (4D PC-MRI). A more in-depth knowledge of these relationships based on 4D PC-MRI has great

potential to improve the diagnosis of CVDs, to assess their severity more reliably

and to support therapeutic decisions through individual risk stratification.

**Data:** From the measured flow information, many attributes can be derived. Besides meta information about the persons, we derived 197 flow attributes. They can be roughly categorized into velocity-related, pressure-related, diameter-related, vortex-related, and flow-jet-related attributes.

* velocity-related: attributes that relate to the speed of the blood flow
* pressure-related: attributes that relate to the pressure of the blood flow
* diameter-related: attributes that relate to the diameter of the aorta
* vortex-related: attributes that relate to blood flow parts that show a vortical behaviour. In a healthy person, the blood flows laminar through the vessel, but in pathological changed aortas the flow starts to swirl inside the vessel. This is called vortical blood flow.
* flow-jet-related: The flow jet describes a part of the flow with highest speed, which is usually centrally located in the aorta. Due to pathological vessel changes the flow jet can be displaced.

The flow information is measured for three classes of persons:

1. Healthy persons
2. Patients with a bicuspid aortic valve (BAV)
3. Patients with a Tetralogy of Fallot (neglect for now, because we have just 10 instances)

**Goal:** We want to evaluate if it is possible to separate the healthy people and BAV patients from each other based on the flow attributes. Many of the attributes correlate with each other. Our goal is to generate an overview visualization, where we can see if both groups can be separated and whether there are any outliers.

**Tasks:**

1. Get familiar with different dimension reduction (DR) techniques (PCA, UMap, t-SNE, multidimensional scaling and maybe other non-linear techniques)
2. Setup an R sheet to load the csv-file
3. Focus the analysis to class 1 and 2 (healthy and BAV)
4. Apply different DR techniques to the attributes in the csv file
5. Visualize the projected results in a scatterplot by given different colors to instances of the two classes
6. Evaluate the separability of the classes after projection (simplest way would be to measure the minimal distance between two instances with different class label), (another way would be to evaluate it qualitatively by generating clusters manually and describe their separability)

**Important Literature:**

* Bloodline: A system for the guided analysis of cardiac 4D PC-MRI data (<http://www.vismd.de/lib/exe/fetch.php?media=files:misc:koehler.pdf>) → with the software Bloodline all the flow attributes of the csv-file were computed. It describes some background of 4D PC-MRI and attribute computation.
* GUCCI - Guided Cardiac Cohort Investigation (paper which is currently under revision. Please do not give it to other people. It describes the current overview visualization using PCA and other aspects of the medical background and attribute computation). It lies in the parent folder of the project.
* UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction (https://arxiv.org/abs/1802.03426)
* G Hinton, S Roweis. “Stochastic neighbor embedding”, *Advances in neural information processing systems* 15, 833-840 (http://papers.nips.cc/paper/2276-stochastic-neighbor-embedding.pdf)
* L van der Maaten, G Hinton. “Visualizing data using t-SNE”, *The Journal of Machine Learning Research* 9 (2579-2605), 85, 2008 (http://www.cs.toronto.edu/~hinton/absps/tsnefinal.pdf)
* M Williams, T Munzner. „Steerable, progressive multidimensional scaling“, *Proc. of IEEE Symposium on Information Visualization*, pp. 57-64, 2004. (https://innovis.cpsc.ucalgary.ca/innovis/uploads/Courses/InformationVisualizationDetails2009/Williams2004.pdf)