Software Engineering Project
Visualization of high-dimensional data

Kiljanek Adrian

Zdziebło Arkadiusz

Stasica Krystian

Ryczek Bartosz

Content

Document history

Version	Description
1.0.0	Basic project concept.
1.0.1	Correction in project description + dictionary

1. Project description

The project is an application to visualize high-dimensional data. Such data can be visualized using: radar plot, parallel coordinates plot. The data consists of parameter's names, the index of objects and numerical values of the parameters of certain objects. The program should support multiple text formats or databases, for example: '.txt', '.xls', '.xlsm', '.csv'. Visualized data can be exported as an image file.

There were no limitations regarding the environment we have to use, thus following the advice of the project supervisor we decided to write our application using python.

We chose python because it is a mature programming language with a mass of libraries that enable comprehensive database analysis. It also has a huge number of tutorials and scientific sources for self-study. Moreover, this is our first project in python, so our team can get to know the new programming language and platform.

2. Dictionary

High-dimensional data - data where number of features or covariates can be larger than the number of independent samples.
Parallel coordinates[1] - way of visualizing high-dimensional datasets. Plot consists of a vertical bar representing a variable. Values are plotted as a series of lines connected across each axis.
Radar chart[2] - also known as spider plot, a chart that consists of a sequence of equi-angular spokes, called radii, with each spoke representing one of the variables.

3. References

[1] https://en.wikipedia.org/wiki/Parallel_coordinates
[2] https://en.wikipedia.org/wiki/Radar_chart