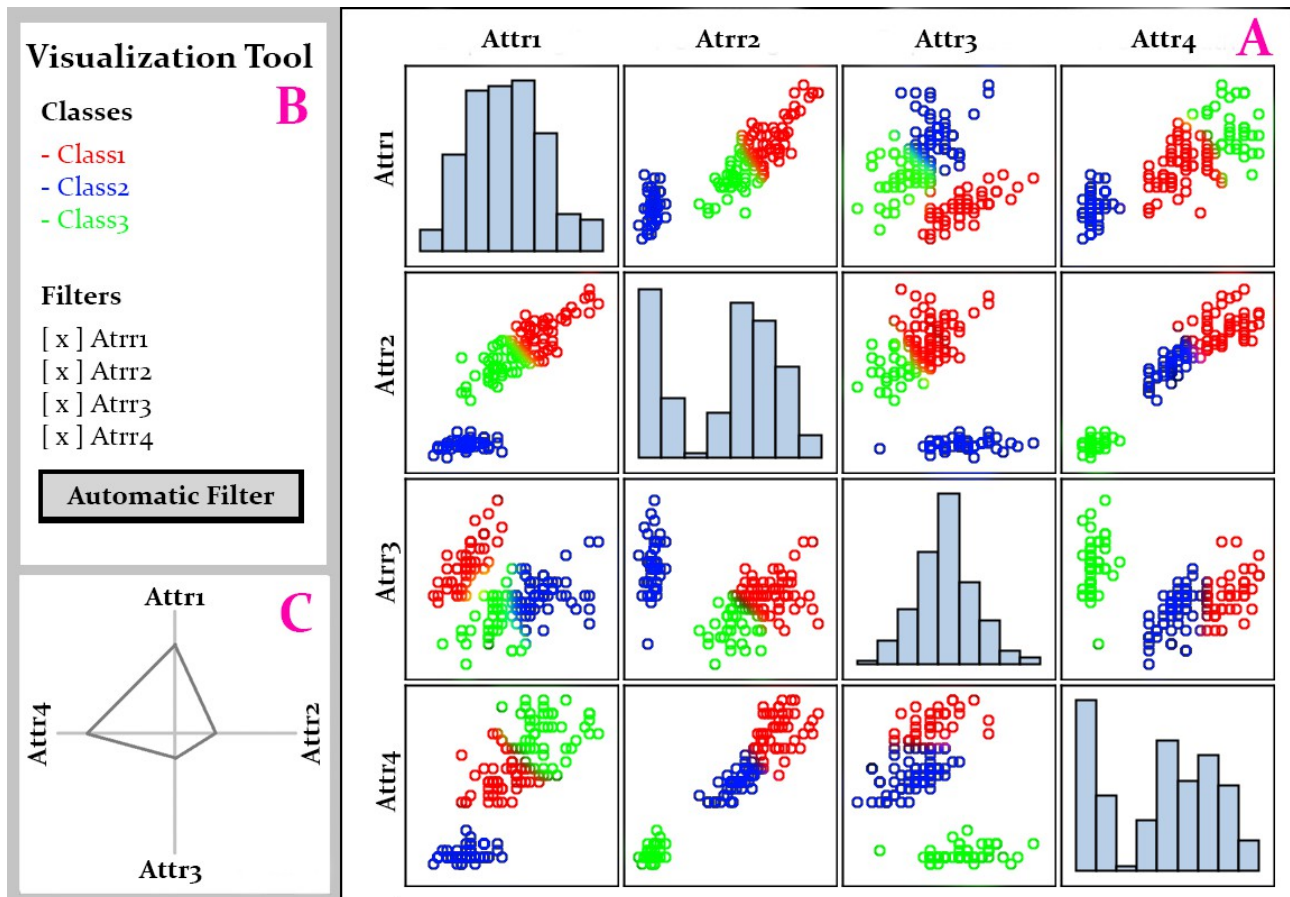


CSCI 6406 Visualization – Project Proposal

Jéssica Pauli de C. Bonson

Goal: Develop a visualization tool for high-dimensional classification data.

Description: The system shows, for a classification real-valued dataset, a scatter-plot matrix for the data attributes, and a diagonal with the histograms for each attribute (part A). The data points are colored depending on their class, and the user can manually or automatically filter the attributes (part B). The automatic filter will use Machine Learning feature filter technique. The user can select data points in a scatter plot, it will highlight the data point in the other scatter plots, and update the star plot (part C) to represent the data point with all the attributes.



Features:

- Selection: Select a point in one of the scatter plots. This and the points in other scatter plots representing this data point will be highlighted, and the star plot will be updated to represent this data point.
- Filter: Add or remove attributes, dynamically changing the scatter plot matrix size and the star plot structure.
- Zoom: Click in a chart to zoom it (scatter or histogram), so it will occupy the whole space instead of the matrix.

Dataset:

- The tool initially will be develop using the [Iris dataset](#), so I can focus more on the tool and less on the complexities of the data. In case there is enough time, the tool will be upgraded to work with data with more features ([seeds dataset](#) (7), [yeast dataset](#) (8), or [breast cancer dataset](#) (32))

Implementation:

- The tool will be implemented using Python and Javascript. I will use the web framework Flask and the Javascript library D3.

In case there is enough time, extra features will be implemented:

- Improve the tool to work with data with more features.
- Automatic Feature Extraction.
- Enable the tool to work with any .csv dataset.

Timeline:

- **9, Feb : Project Proposals**
- **14, Feb:** Implement the general structure and organization of the system, in a way so that the system parts and features can be developed and integrated cleanly.
- **2, Mar:** Completed part A.
- **9, Mar: Project Updates:** Completed parts B and C, but with the three features (selection, filter, and zoom) not working yet.
- **16, Mar:** Implemented the selection feature
- **29, Mar:** Implemented the filter feature (manual and automatic)
- **5, Apr:** Implemented the zoom feature. The next days are to solve bugs, improve the tool, or implement extra features.
- **9, Apr: Final Projects:** Three system parts working and three features working.