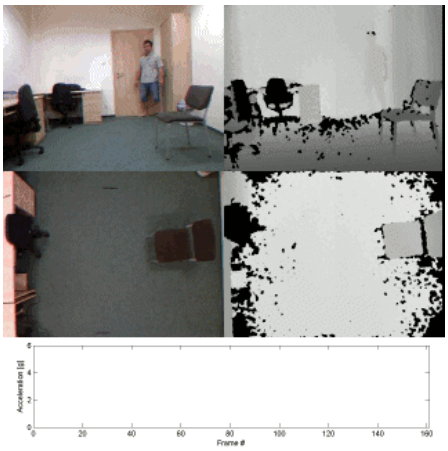
Project 1

UR Fall Detection Dataset



http://fenix.ur.edu.pl/~mkepski/ds/uf.html

Data Project Questions & Considerations

PACE: Plan Stage

- What are the data columns and variables and which ones are most relevant to your delivered

The dataset are csv files with data from an accelerometer. Each csv file contains a number for total, x, y and z, for both: fall and not-fall tests.

- What units are your variables in?

All variables are ‘float’.

- What are your initial presumptions about the data that can inform your EDA, knowing you will need to confirm or deny with your future findings?

First of all, it is necessary to conduct an individual study of the data by creating graphics for fall and non-fall data to better visualize the information.

- Is there any missing or incomplete data?

No. there is not any missing data.

- Are all pieces of this dataset in the same format?

Yes, the dataset consists solely of numerical values (floats).

- Which EDA practices will be required to begin this project?

Data collection, Visualizing and Analyzing Results with graphics.

PACE: Analyze Stage

- What steps need to be taken to perform EDA in the most effective way to achieve the project goal?

Code a program to extract the CSV files from the webpage. Compile 10 graphics for the fall and 10 for the not-fall CSV files. Calculate the average, maximum, and minimum for each axis of the accelerometer and create a matrix with these numbers.

- Do you need to add more data using the EDA practice of joining? What type of structuring needs to be done to this dataset, such as filtering, sorting, etc.?

I will add a new dataframe with the calculated values.

- What initial assumptions do you have about the types of visualizations that might best be suited for the intended audience?

The most applicable graphic in this case is the line graph.

PACE: Construct Stage

- What data visualizations, machine learning algorithms, or other data outputs will need to be built in order to complete the project goals?

To achieve the project goals, it will be necessary to use the perceptron algorithm.

- Which variables are most applicable for the visualizations in this data project?

The average, maximum, and minimum for each axis of the accelerometer.

- Going back to the Plan stage, how do you plan to deal with the missing data (if any)?

There is not missing data.

PACE: Execute Stage

- What key insights emerged from your EDA and visualizations(s)?

After obtaining the graphics, I need to analyze each one and understand where the biggest differences in numbers lie. I will then use these most discrepant numbers between the fall and not-fall graphics to build my array of selected features. This approach will help me enhance the outcomes of the perceptron algorithm.

- What business and/or organizational recommendations do you propose based on the visualization(s) built?

not applicable

- Given what you know about the data and the visualizations you were using, what other questions could you research for the team?

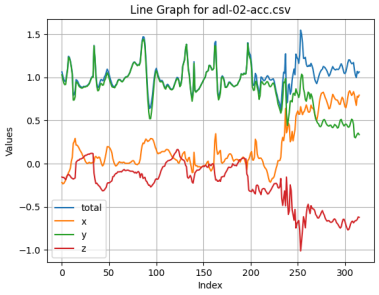
not applicable

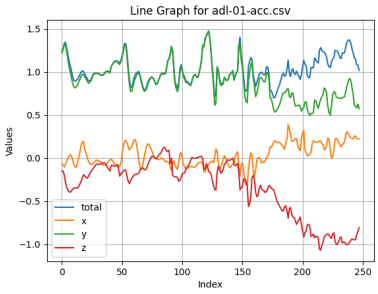
- How might you share these visualizations with different audiences?

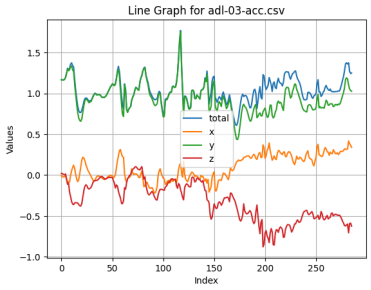
For better visualization, I might share the graphics and the perceptron's plots, which show the outcomes more clearly

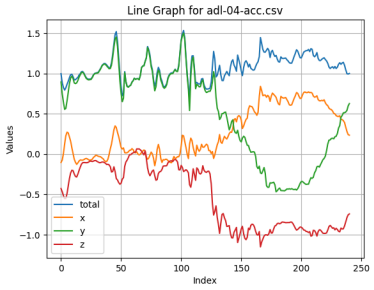
Data visualization

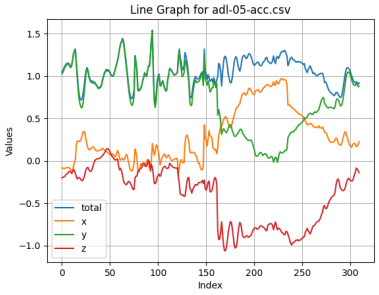
Graph for adl

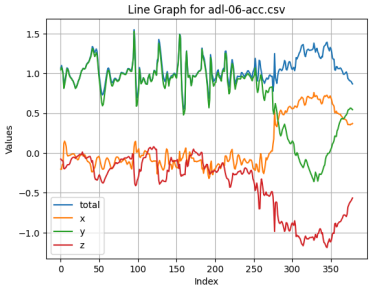


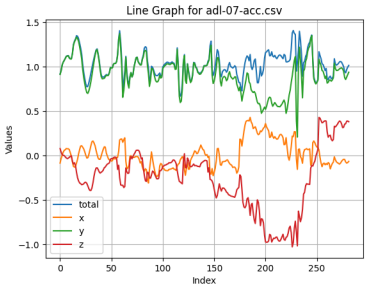


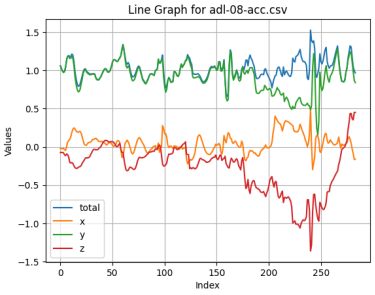


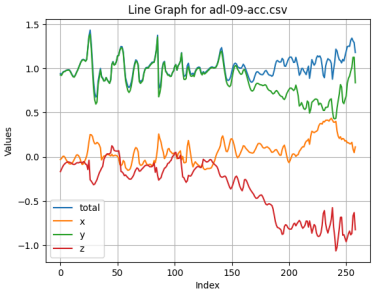




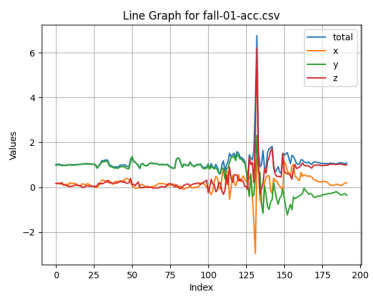


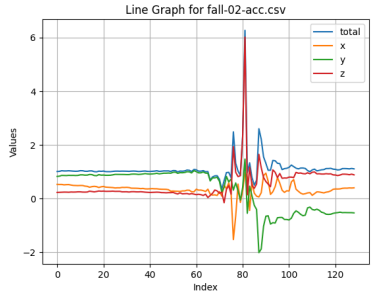


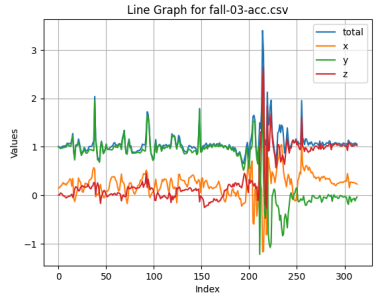


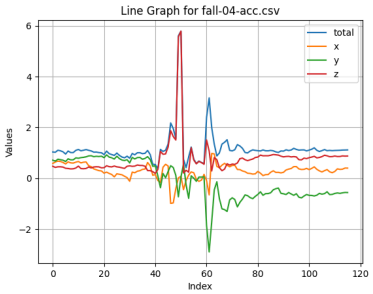


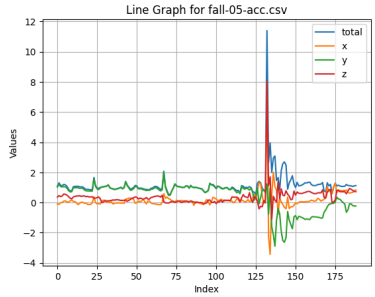
Graph for fall

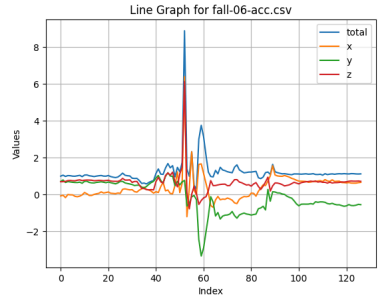


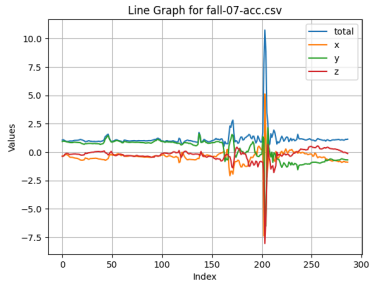


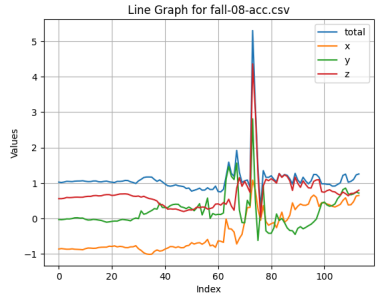


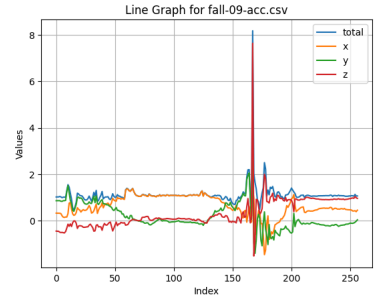






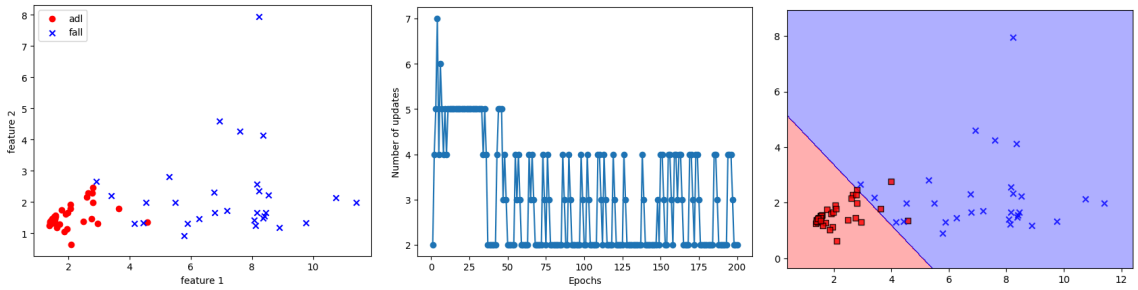




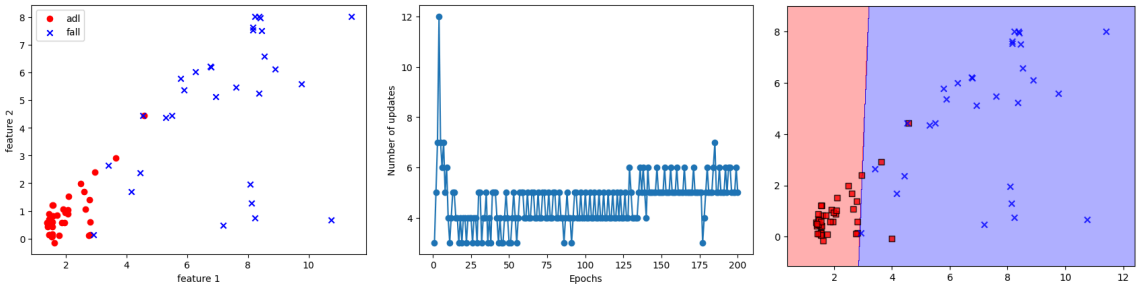


Model Evaluation and Validation

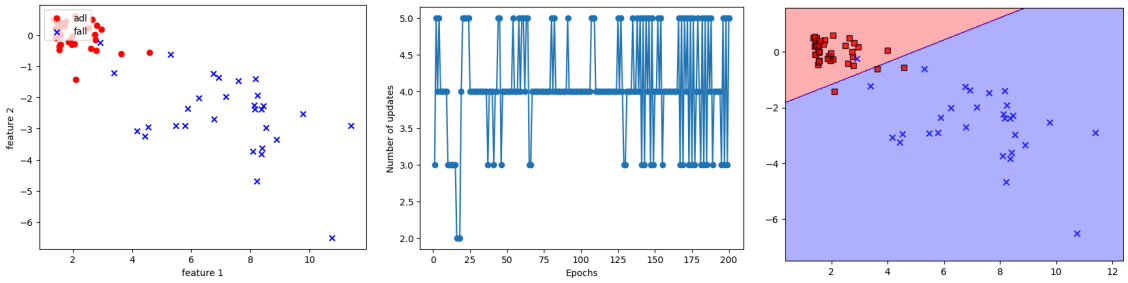
Features: ‘means_total’ and ‘max_y’



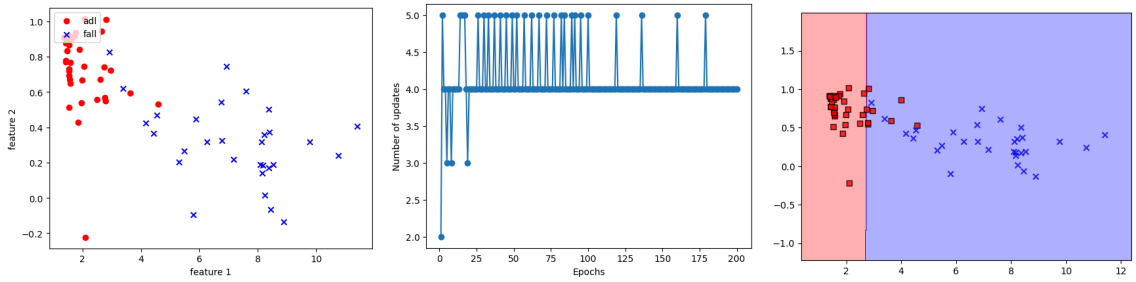
Features: ‘max_t’ and ‘max_z’



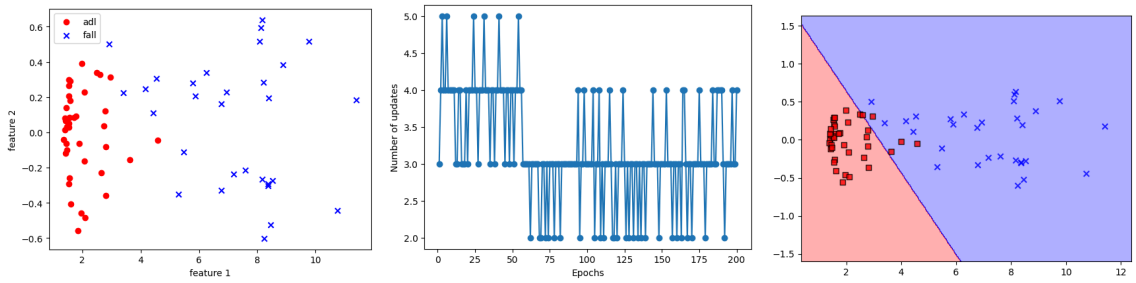
Features: ‘max_t’ and ‘min_y’



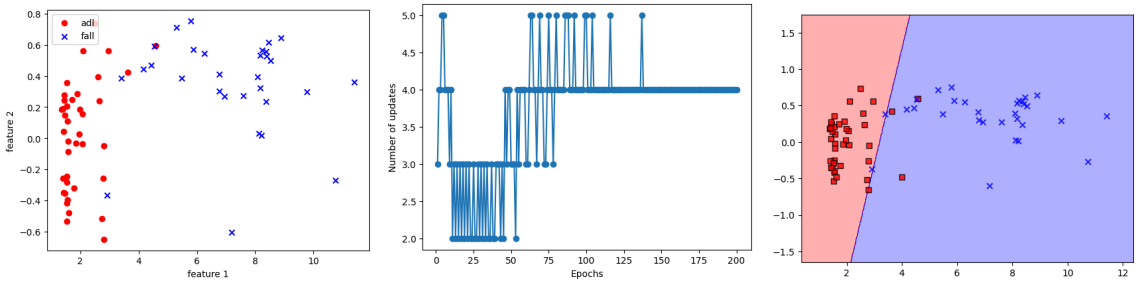
Features: ‘means_y’ and ‘max_t’



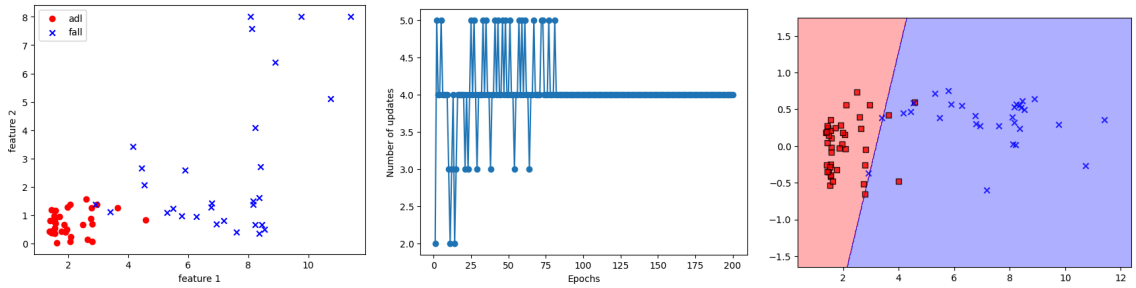
Features: ‘max_x’ and ‘max_t’



Features: ‘means_z’ and ‘max_t’



Features: ‘max_t’ and ‘max_x’



Features: ‘max_t’ and ‘min_y’