Data visualization course Final Project

Student name: Lior Breitman

I.D: 212733257

During the semester, the 2 main things we focused on were ensemble learning and dimensionality reduction.

Our project at the end had 4 jupyter notebooks, but I will explain them in 2 parts, the first 3 notebooks, and the last one.

Part one:

In the first 3 notebooks, we had the classification notebook from last semester, as well as one notebook where we worked on the Fashion-Mnist dataset, and the third where we worked on the Cats-Vs-Dogs data set.

In all of those notebooks, the data was clean, the first notebook I cleaned during last semesters project, and the later 2 because they are very famous data sets.

In the last notebook I had to learn how to import images and change them to arrays.

Then I used various methods of ensemble learning:

- Voting, both hard and soft
- Bagging
- Boosting
- Stacking

I also used the PCA algorithm to try and see how much dimensions are actually needed to explain a certain amount of the Variance.

Part two:

In the second part, we got a bunch of data files which were recordings of peoples hand movement in an experiment. They were split to 3 types, when the person was alone, when the person was in front of another one and tried to sync up with their movement and again in front of another person, but this time trying to be spontaneous.

In this part I had to import all the files, and merge them together according to the right and left hand recordings of the second frame, and merge the alone files with a specific file.

After this I showed some of the movements in plots to see how they compare.

Then I again started training models and using ensemble methods to try and get the best accuracy I can for this data.