Capstone Project Proposal

Name: Eva Schwartz

Objective

Using biological health data such as heart rate and skin temperature derived from wrist-worn wearable devices to detect physiological stress, such as illness or injury, which can be indicated through anomalies in heart rate variation or respiration rate.

Impact

This project has potential impacts in the wearable devices market, while demonstrating ways that consumers/users of health data can take their health into their own hands using data they are already collecting. Part of my project will be researching how this could be implemented onto a wearable device using TinyML, a machine learning framework that focuses on deployment on smaller devices like smartwatches.

Dataset(s)

My datasets are uploaded to my repository on Github under the data folder. The training dataset is for training, testing, and validating the various stress classification models, then the testing dataset is for testing the real-life deployment of my chosen model/models on data that was collected on the same type of device, but in potentially different settings, representing the real deployment methods for developing software.

Both the training and testing datasets are reliable and have been used in many other settings.

Approach

1) Finalize goals/process for the project
2) Work on preprocessing data, and understanding when/if I will need to denoise
3) Data cleaning/preparation
4) Model training and testing on training data
5) Model deployment on testing data
6) Bigger picture – so what? Of the project
7) Presentation – I really want to have a strong presentation, implementing project management ideas throughout my capstone ☺
8) Write research report discussing idea and methods, then make poster

Timeline – dates represent internal due dates for each part

1) Finalize goals/process for the project/implement professor feedback on capstone + ideas (January 24)
2) Work on preprocessing data, and understanding when/if I will need to denoise (February 7)
3) Data cleaning/preparation (February 28)
4) Model training and testing on training data (March 14)
5) Model deployment on testing data (March 28)
6) Bigger picture – so what? Of the project (April 4)
7) Summarizing my project – presentation, write research report discussing idea and methods, then make poster (April 18)

Possible Issues

- Splitting time series data – I identified this as an issue from the beginning, and have done background research on how to navigate this ☺
- Making sure my question is complex enough – I feel like the data processing/understanding the data is one of the bigger challenges of this project, which I am excited about, but I also want to make sure my question is complex enough which I am a little bit concerned about
    - I know that in this class, I am limited by the available data, and while there might be questions that I am interested in examining, the data might not always be there, so this is a good start
- Understanding whether I want to do one research paper or two – I feel like this is something that I will kind of sort out along the way…. If I am seeing a big business thread/idea in my project then maybe I will write one focused on data science/ML and one on the business perspectives from implementing software like this but I don't think I will know this until I keep working with the data