**Stress-Predict Dataset**

Ensure that the folder containing the dataset is inside the same path as the Jupiter notebook.

The Jupiter notebook lists all dependencies in the start, ensure these are installed within the environment before running the notebook.

use either pip or Conda to install before running if any are missing.

The notebook contains clear sections to demonstrate what is happening, all paths are written to be independent of the global path if the notebook is in the same folder as the folder containing the data.

All code is self-contained within the notebook and does not have any external dependencies other than libraries to be installed in the environment.

**Dataset Files**

**|-- *Processed\_data***

|----- ***heartrate\_resprate\_timestamps\_labels folder***

|----- ***Improved\_Combined\_hr\_rsp\_binary\_PX.csv*** (contain information of heart rates and respiratory rates along with timestamps and labels (for nonstress/baseline and 1 for stress task duration). Here *X* is participant number)

|----- ***Time\_logs.xlsx*** (contain date and start/end time of each task for each participant, Irish standard time)

|----- ***heartrate\_timestamps\_labels folder***

|----- ***PX\_comb\_binary.csv*** (contain information of heart rates along with timestamps and labels (for nonstress/baseline and 1 for stress task duration). Here *X* is participant number)

|----- ***resprrate\_timestamps\_labels folder***

|----- ***Improved\_PX\_comb\_10sec\_binary.csv*** (contain information of respiratory rates along with timestamps and labels (for nonstress/baseline and 1 for stress task duration). Here *X* is participant number)

|----- ***Improved\_All\_Combined\_hr\_rsp\_binary.csv*** (contain information of heart rates and respiratory rates of all the participants along with timestamps and labels (for nonstress/baseline and 1 for stress task duration))

|----- ***Questionnaires\_scores.xlsx*** (contains information about the PSS and STAI questionnaire scores of each participant)

|----- ***Time\_logs.xlsx*** (contain date and start/end time of each task for each participant, Irish standard time)

**|-- *Raw\_data***

|----- **S*X* *folder*** (folders with raw files from Empatica E4. Where *X* is participant number)

|----- ***ACC.csv*** (contains accelerometer data (x, y, z axis))

|----- ***BVP.csv*** (contains raw BVP data)

|----- ***EDA.csv*** (contains EDA data (skin conductance))

|----- ***HR.csv*** (contains heart rate data)

|----- ***IBI.csv*** (contains inter-beat-interval data)

|----- ***info.txt*** (contains information about all the csv file and sampling rate)

|----- ***tags\_SX.csv*** (contains timestamp tags. start-end time of each task)

|----- ***TEMP.csv*** (contains skin temperature data)

These all files are used from original dataset

**Libraries**

Following libraries were used for analysis:

* Statistical Analysis (Rstudio):
  + Numpy
  + Pandas
  + Matplotlib
* Descriptive Analysis (python):
  + Pandas
  + Numpy
  + Matplotlin
* Classification Analysis (python):
* Scikit-learn
* Decision Tree
* Gradient Boost Classifier
* Randomized Search

**References**

When citing this dataset, please also use following citation:

1. N. C. Basjaruddin, F. Syahbarudin, and E. Sutjiredjeki. Measurement device for stress level and vital sign based on sensor fusion. Healthcare Informatics Research, 27(1):11–18, 2021.
2. Y. S. Can, N. Chalabianloo, D. Ekiz, and C. Ersoy. Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study. Sensors, 19(8):1849, 2019.
3. T. Iqbal, A. J. Simpkin, D. Roshan, N. Glynn, J. Killilea, J. Walsh, G. Molloy, S. Ganly, H. Ryman, E. Coen, et al. Stress monitoring using wearable sensors: A pilot study and stresspredict dataset. Sensors, 22(21):8135, 2022.
4. B. Szakonyi, I. Vass´anyi, E. Schumacher, and I. K´osa. Efficient methods for acute stress detection using heart rate variability data from ambient assisted living sensors. BioMedical Engineering OnLine, 20(1):1–19, 2021.