# Stress level detector

## About the data

Train data and test data each has 3 files:

1. Frequency domain features
2. Heart rate non linear features
3. Time domain features

Combined the 3 files into one as train.csv and test.csv respectively.

Total columns: 37 (excluding the common ones)

Only considered

features = ['MEAN\_RR','RMSSD','pNN25','pNN50','LF','HF','LF\_HF']

Mean RR refers to the **mean value of the RR intervals** in a heart rate variability (HRV) analysis. An RR interval is the time interval between two successive R-wave peaks in the electrocardiogram (ECG).

**At Rest**: 600 ms to 1200 ms

RMSSD (Root Mean Square of Successive Differences) is a time-domain measure of Heart Rate Variability (HRV). It quantifies the short-term variability in heart rate by calculating the root mean square of the differences between successive RR intervals (time between heartbeats).

Low RMSSD (<20 ms) is associated with stress, anxiety, or poor autonomic regulation.

 **Normal RMSSD**: 20–50 ms

 **Highly Fit Individuals**: RMSSD can exceed 100 ms

pNN25 (Percentage of NN intervals differing by more than 25 milliseconds) is a time-domain measure of Heart Rate Variability (HRV). It represents the percentage of consecutive RR intervals (NN intervals) that differ by more than 25 milliseconds. It is another metric that reflects the parasympathetic (vagal) activity of the autonomic nervous system.

**Healthy Adults**:

* **Normal pNN25**: Typically **3–10%**.
* Higher percentages indicate better vagal tone and relaxation.

**Children and Young Adults**:

* pNN25 is higher, as they tend to have greater heart rate variability.
* Ranges may exceed **10–20%**.

**Stress or Illness**:

* A lower pNN25 (<3%) may indicate reduced parasympathetic activity or increased sympathetic dominance (e.g., stress, fatigue, cardiovascular issues).

pNN50 (Percentage of NN intervals differing by more than 50 milliseconds) is a time-domain measure of Heart Rate Variability (HRV). It calculates the percentage of adjacent RR intervals (or NN intervals) that differ by more than 50 milliseconds. Like pNN25, it reflects **parasympathetic (vagal) activity** of the autonomic nervous system and is often used as an indicator of cardiac health and stress levels.

**Healthy Adults**:

* **Normal pNN50**: **>20%** is considered healthy, though values may vary between **10–40%**.

**Children and Young Adults**:

* **Higher pNN50** values, often exceeding **30–40%**, are common due to greater parasympathetic activity.

**Stress or Illness**:

* **Lower pNN50** (<10%) may indicate **reduced vagal tone**, increased **sympathetic activity**, or overall poor cardiovascular adaptability (e.g., due to chronic stress, fatigue, or medical conditions).

LF (Low-Frequency power) is a frequency-domain measure of **Heart Rate Variability (HRV)**. It represents the power spectral density of heart rate fluctuations in the **low-frequency band**, typically ranging from **0.04 to 0.15**

HF (High-Frequency power) is a frequency-domain measure of **Heart Rate Variability (HRV)**. It represents the power spectral density of heart rate fluctuations in the **high-frequency band**, typically ranging from **0.15 to 0.4**

LF/HF ratio is a commonly used metric in heart rate variability (HRV) analysis. It represents the balance between **Low-Frequency (LF)** power and **High-Frequency (HF)** power. This ratio provides insight into the balance between the **sympathetic** (fight-or-flight) and **parasympathetic** (rest-and-digest) branches of the autonomic nervous system (ANS).

A **high LF/HF ratio** suggests **sympathetic dominance** (stress or arousal).

## Errors encountered

The combined files were not properly included all the columns when I tried to do it with concat. Used merge function to solve the error.

Error encountered: used the to\_categorical function which requires numeric values for the labels, but y still contained string values like 'no stress'.

So converted those string labels into integers:

1. **Convert String Labels to Integer Labels**: Used LabelEncoder from sklearn to map string labels like 'no stress' to integer labels.
2. **Apply to\_categorical**: Once the labels are integers, converted them to one-hot encoding.