

(Photo by [Jared Rice](https://unsplash.com/@jareddrice?utm_source=unsplash&utm_medium=referral&utm_content=creditCopyText) on [Unsplash](https://unsplash.com/photos/NTyBbu66_SI?utm_source=unsplash&utm_medium=referral&utm_content=creditCopyText))

PRD: Stress Prediction System

# Problem Alignment

## The Problem

The opportunity comes when stress can lead to serious illness in the future if not properly taken. In journal “The effects of chronic stress on health: new insights into the molecular mechanisms of brain–body communication” it shows that Diseases whose development has been linked to both stress and inflammation include cardiovascular dysfunctions, diabetes, cancer, autoimmune syndromes and mental illnesses such as depression and anxiety disorders. So it will be better if we can detect stress levels in our body, therefore we can take precaution early before serious illness comes.

By addressing the challenge, we create a system that can predict stress levels based on metric healthy lifestyle. Hence the people will take precautionary action early.

## High-Level Approach

## Stress level prediction will be using machine learning methodology and healthy lifestyle features in the dataset to decide stress level on a certain person. The system will show the stress level of a person after getting info of some healthy lifestyle input such as sleep disorder, blood pressure, BMI, etc.

## Goals & Success

## I will use metric level stress and translate it into ML metrics with 3 different metrics (MAPE, RMSE, and R2). These 3 ML Metrics would represent the success of the model and predict “level stress” from the user. The reason for choosing 3 ML metrics, because I need to ensure accuracy of a model that predicts the stress level. By comparing those 3 metrics, I might deliver a more trusted result in terms of prediction.

# Solution Alignment

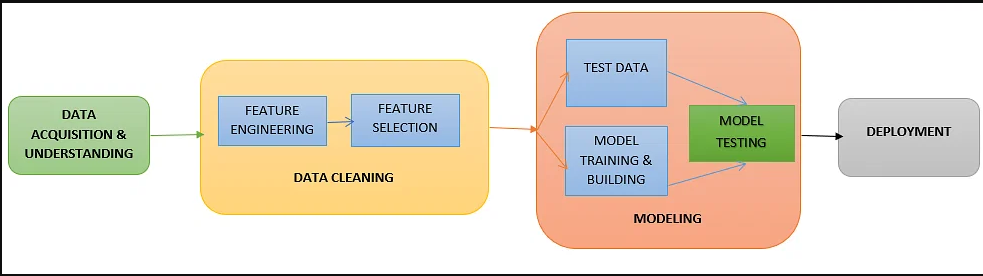
## Key Solution

In this project we will be using <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset> as a dataset. The dataset consists of 400 rows and 13 columns, containing various variables related to sleep health and lifestyle. The columns and their descriptions are as follows:

1. Person ID: Unique identifier for each individual.
2. Gender: Gender of the person (Male/Female).
3. Age: Age of the person in years.
4. Occupation: Person's occupation or profession.
5. Sleep Duration (hours): Number of hours slept by the person in a day.
6. Sleep Quality (scale: 1-10): Subjective evaluation of sleep quality on a scale from 1 to 10.
7. Physical Activity Level (minutes/day): Number of minutes spent on daily physical activity.
8. **Stress Level (scale: 1-10): Subjective evaluation of stress level on a scale from 1 to 10.**
9. BMI Category: BMI category of the person (e.g., Underweight, Normal, Overweight).
10. Blood Pressure (systolic/diastolic): Measurement of blood pressure represented through diastolic over systolic pressure.
11. Resting Heart Rate (bpm): Resting heart rate of the person in beats per minute.
12. Daily Steps: Number of steps taken by the person in a day.
13. Sleep Disorder: Presence or absence of a sleep disorder in the person (None, Insomnia, Sleep Apnea).

I will use Linear Regression as the main model methodology for the system.

## Key Flows



# 

# Launch Readiness

## Key Milestones

| **Date** | **Milestone** | **Description** |
| --- | --- | --- |
| Sept 13 | Brainstorm Idea | Brainstorm idea and what are problem need to be solve |
| Sept 15 | Gathering Dataset | Finding dataset |
| Sept 18 | PRD docs | Finish Creating PRD Docs |
| Sept 20 | EDA Stage 1 | Early EDA, gathering info data, making some question for stage 2 EDA |
| Sept 23 | EDA Stage 2 | Advance EDA, conducting some AB Test, answering some question |
| Sept 21 | Feature Engineering | Performing feature engineering numerical and categorical feature |
| Sept 25 | Feature Selection and split dataset | Choosing some feature and conduct splitting dataset into train and test |
| Sept 30 | Modeling | Create ML model and conduct train & test, hyperparameter tuning |
| Oct 5 | Deploying | Deploy model ML into UI web |
| Oct 12 | Reporting | Finish Creating Final Report |
| Oct 13 | Video Demo | Finish Creating video demo |

## Artifacts

*Put all the details and models of your milestone*

| **Artifacts** | **Where to check?** |
| --- | --- |
| Dataset Final | https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset |
| Github | https://github.com/WidharDwiatmoko/stress-level-prediction |
| App | TBU |

## References

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5137920/#:~:text=Diseases%20whose%20development%20has%20been,as%20depression%20and%20anxiety%20disorders>.
2. Kaggle Dataset : Sleep Health and Lifestyle Dataset
3. Pacmann Course : Intro ML and ML Process