**Sleep Disorder Prediction**

**Objective**

To predict whether a person has a sleep disorder, i.e., sleep apnea or insomnia. The problem addressed is aligned with the UN SDG 3 Health.

**Data Source**  
The dataset used was obtained from Kaggle: <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>

**Overview**

The dataset consists of 374 rows and 12 columns with a wide range of variables related to sleep and daily habits. They include gender, age, occupation, sleep duration, quality of sleep, physical activity level, stress levels, BMI category, blood pressure, heart rate, daily steps, and the presence or absence of sleep disorders. The target variable is **the presence or absence of a sleep disorder.**

**Feature Categorization**

**Comprehensive Sleep Metrics**: Explore sleep duration, quality, and factors influencing sleep patterns.

**Lifestyle Factors**: Analyze physical activity levels, stress levels, and BMI categories.

**Cardiovascular Health**: Examine blood pressure and heart rate measurements.

**Sleep Disorder Analysis**: Identify the occurrence of sleep disorders such as Insomnia and Sleep Apnea.

**Machine Learning Approach**

This is a Classification Problem where the input features were used to categorize a person’s sleep disorder. Various classification models were used. They include Support Vector Machine, Decision Trees, and Random Forest. The metrics used were precision, recall, and the F1-score. Accuracy was not the best evaluation metric since the data was imbalanced. Random Forest performed well, followed by Decision Trees and Support Vector Machine.

**Observation**

There was a strong correlation between quality of sleep and sleep duration, daily steps and physical activity levels, and stress level and heart rate. Additionally, there was a weak correlation between age and quality of sleep. The data was obtained from people of ages ranging from 27 to 59.