

Project Proposal

Title: Predicting Sleep Disorders Based on Lifestyle and Health Data

Team Details

Team Name: Three-Eyed Ravens

Team Members:

Amishi Patidar (24BCS10184), Suhani Awasthi (24BCS10260), Khushboo Sigroha (24BCS10340)

Problem Statement

Sleep disorders such as insomnia, sleep apnea, and poor sleep quality are increasingly prevalent due to modern lifestyle and health factors. They negatively impact physical health, mental well-being, productivity, and overall quality of life. Early prediction of potential sleep disorders can prevent escalation, reduce healthcare costs and encourage healthier routines.

Data Source

We will use the **Sleep Health and Lifestyle Dataset** from Kaggle:

<https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>

Some key features include: **Demographics**: Age, Gender, Occupation, **Lifestyle**: Work hours, Exercise, Screen time, Caffeine/Alcohol intake, **Health Indicators**: BMI, Blood Pressure, Stress levels, **Sleep Data**: sleep duration, sleep quality, sleep disorder diagnosis.

Approach

1. **Data Preprocessing**: Handle missing values, encode categorical variables, normalize continuous data, and manage outliers.
2. **Exploratory Data Analysis (EDA)**: Visualise correlations between lifestyle & sleep, and identify patterns (e.g., stress/work hours impact).
3. **Model Development**: Apply Logistic Regression (disorder classification), and Linear Regression (sleep duration/quality).
4. **Model Evaluation**: Accuracy, Precision, Recall, F1, ROC-AUC (classification); RMSE, MAE (regression).

Expected Outcomes

- A predictive model to estimate risk of sleep disorders based on health and lifestyle factors.
- Provide insights into major factors affecting sleep quality and disorder likelihood, enabling targeted health interventions.
- A framework extendable to preventive healthcare tools in workplaces, universities, and clinics.
- Increased awareness of how lifestyle choices affect long-term sleep health.