ECS 171 Group Project

Amira Basyouni, Alexis Lydon, Calvin Chen, Ryan Yu, Tianming Tan

October 14, 2023

Problem Statement

Sleep, as a student, many of us don't get enough of it. We've heard it affects our health, but in what ways? How much is enough sleep and why should we care? After analyzing our data, we can hopefully understand why this is important and improve our sleep quality. We will be able to narrow down the factors that affect our sleep, and build habits that can mitigate and reduce the negative effects sleep deprivation has on our health. Through the use of Machine Learning models, we can create predictions on what factors would lead to better or worse sleep, and predict if a person would have high or low sleep quality based on these factors. Not only will this help students like us better manage our sleep, but it will also help us as a community of diverse occupations to get a good night's rest.

Dataset

The dataset being used in this project is the Sleep Health and Lifestyle dataset by Laksika Tharmalingam on Kaggle. The dataset comes in the form of a CSV file and contains 400 rows and 13 columns, encompassing various sleep and lifestyle variables. These include gender, age, occupation, sleep duration, sleep quality, physical activity, stress levels, BMI category, blood pressure, heart rate, daily steps, and sleep disorder status. We chose this dataset because, as students, we can relate to the common sleep issues many of us face. Our project focuses on investigating the relationship between sleep and health, making a dataset on sleep and health the most suitable choice. A limitation of the dataset is its synthetic nature, generated artificially rather than from real-world observations. While it may lack some real-world nuances, high-quality synthetic data can effectively train and test machine learning models in this context.

1 Goals

Project Timeline

For our project, we plan to roughly follow the timeline below:

Project Timeline			
Task	**Start Date**	**End Date**	**Duration (days)**
Background/Literature Review	Oct. 16	Oct 22.	7
Exploratory Data Analysis	Oct 23.	Oct 29.	7
Developing prediction models	Oct 30.	Nov. 5-8	7-10
Evaluation of the model and testing performance	Nov. 6-9	Nov. 12-14	7-9
Developing front-end to display and run models	Nov. 13-15	Nov. 19-21	7
Finish Report and Minor Adjustments	Nov. 20-22	Nov. 23-25	4
Presentation Practice	Nov. 24-26	Nov. 28-30	5

Table 1: Tasks, start and end dates, and planned duration of each task.