

Workshop 2: System Design for Detecting Sleep States Using Accelerometers

Juan Carlos Quintero Rubiano

Code: 20232020172

Systems Engineering

Francisco Jose de Caldas District University

Bogota, Colombia

jcquineror@udistrital.edu.co

Juan Nicolas Diaz Salamanca

Code: 20232020059

Systems Engineering

Francisco Jose de Caldas District University

Bogota, Colombia

jndiazs@udistrital.edu.co

Juan Felipe Wilches Gomez

Code: 20231020137

Systems Engineering

Francisco Jose de Caldas District University

Bogota, Colombia

jfwilchesg@udistrital.edu.co

Abstract—This document builds upon the systemic analysis conducted in Workshop 1 for the Kaggle competition "Child Mind Institute - Detect Sleep States." The focus is on designing a robust system to detect sleep states using accelerometer data. Key findings from the previous analysis, including constraints, data characteristics, and chaos-theory factors, are summarized to guide the design process. The proposed design aims to address identified weaknesses and optimize system performance.

Index Terms—System design, sleep state detection, accelerometers, chaos theory, optimization.

I. INTRODUCTION

A. Overview of Workshop 1 Findings

The systemic analysis conducted in Workshop 1 provided a comprehensive understanding of the system for detecting sleep states using accelerometer data. The following key findings were identified:

- **Constraints:**
 - Sleep periods must be at least 30 minutes long, with interruptions no longer than 30 minutes.
 - Only the longest sleep window per night is recorded.
 - Device removal periods are not annotated, introducing potential gaps in data.
- **Data Characteristics:**
 - The dataset includes accelerometer data with features such as `anglez` and `enmo`, which are critical for detecting sleep states, and a `timestamp` column for time-based analysis.
 - Labels for sleep onset and wake events are provided, enabling supervised learning approaches.
- **Chaos-Theory Factors:**
 - Sensitivity to initial conditions: Small variations in accelerometer data can lead to significant changes in sleep state classification.

- Randomness in sleep patterns: External factors such as environmental conditions and individual differences introduce unpredictability.

B. Insights from Systemic Analysis

The systemic analysis provided the following critical insights for designing the system:

- **Feature Importance:**
 - The accelerometer features `anglez` and `enmo` are pivotal for detecting sleep states. Proper preprocessing, such as normalization and noise filtering, is essential to maximize their utility.
 - The `timestamp` column enables time-series analysis, which is crucial for capturing temporal dependencies in sleep patterns.
- **Handling Data Gaps:**
 - Device removal periods and unannotated gaps in the data introduce missing values that require robust imputation techniques or strategies to handle incomplete data effectively.
 - Strategies such as forward-filling, interpolation, or machine learning-based imputation can be explored to address these gaps.
- **Temporal Dependencies:**
 - Sleep states are inherently sequential and time-dependent. Leveraging models designed for time-series data, such as Long Short-Term Memory (LSTM) networks or Temporal Convolutional Networks (TCNs), can improve classification accuracy.
 - Capturing transitions between sleep and wake states is critical for accurate predictions.
- **Model Robustness:**

- The system must account for sensitivity to initial conditions, as small variations in accelerometer data can lead to significant changes in classification outcomes.
- Randomness in sleep patterns, influenced by external factors such as environmental conditions or individual variability, requires the model to generalize well across diverse scenarios.
- Techniques such as data augmentation, ensemble modeling, and robust validation can help improve model reliability.

- **Optimization Goals:**

- The design should focus on minimizing false positives (e.g., misclassifying wake states as sleep) and false negatives (e.g., missing sleep onset events) to ensure high precision and recall.
- Computational efficiency is critical, especially if the system is intended for real-time or large-scale deployment.