**Introduction**

The goal of this project is to utilize a Large Language Model (LLM) to reproduce a research paper focused on computational healthcare. The target paper, “ Bringing At-home Pediatric Sleep Apnea Testing Closer to Reality: A Multi-modal Transformer Approach” (Fayyaz et al.. 2023), aimed to explore potential computational tools for use in the diagnosis of pediatric obstructive sleep apnea/hypo-apnea syndrome (OSAHS).

Pediatric obstructive sleep apnea is a sleep-related disorder in which periods of intermittent obstruction of the upper airway occur during a child’s sleep, which can lead to significant adverse health effects. (Gupta et al .. 2024). This disorder can affect ~ 1-5% of children between ages 5 and 8. The authors of the original paper identified a gap in studies performed on pediatric OSAHS, despite a significant amount of literature on adult sleep apnea. The aim of the original method was to continue to expand the efficacy and potential for diagnostic methods for pediatric OSAHS, increasing the availability of accessible diagnostic methods for providers to help children with this condition.

**Data Access and Implementation Details**

The original paper utilized two datasets, the Nationwide Children’s Hospital Sleep DataBank (NCH), and the Childhood Adenotonsillectomy Trial (CHAT) datasets. These datasets supply the necessary polysomnography data as well as additional characteristic data for children age 0-6 years. Access to this data is requested through the governing bodies which manage the datasets (National Sleep Research Resource [NSRR] and BioLINCC respectively). The authors also maintained a repository for the developed codebase3 which will be used to aid in the recreation in this project. This repository contains all code used to generate the model used in the paper, which will be used to recreate the results. Additionally, the code used to preprocess the datasets will be identical in this recreation in order to maintain the same conditions of the original computations.

As shown in the original report, there are approximately 4000 patients on which sleep studies were performed. With the modalities of data collected (EEG, EOG, ECG, SpO2, ETCO2), there is a significant amount of data for which to process. Every attempt will be made to utilize the same portion of the dataset as the original paper, though a subset of the data will be taken if necessary to mitigate extreme computation time.

**REFERENCES**

1. Fayyaz. Hamed, Strang. Abigail, Beheshti, Rahmatollah, “Bringing At-home Pediatric Sleep Apnea Testing Closer to Reality: A Multi-modal Transformer Approach” *Proc Mach Learn Res.* August, 2023 [Online]

<https://pmc.ncbi.nlm.nih.gov/articles/PMC10854997/pdf/nihms-1925296.pdf>

1. Gupta. Surendra, Sharma. Rakesh, “Pediatric Obstructive Sleep Apnea: Diagnostic Challenges and Management Strategies”, *Cureus* 16(12): e75347. Doi:10.7759/cureus.75347
2. Fayyaz. Hamed, “Pediatric-Apnea-Detection”, GitHub, 2023 [Online] Available: https://github.com/healthylaife/Pediatric-Apnea-Detection