Nordic AI Meet 2026

O6. AI for Science & Engineering (Thu 11:15-12:15)

3. TRUST-Sleep: A Trustworthy Event-Centric Framework for Sleep Apnea Screening with Uncertainty-Aware Clinical Decision Support.

Authors: Poushali Sengupta1; Christopher Frenning2; Britt Øverland3; Vera Hermine Goebel1; Thomas Peter Plagemann1

Affiliations: 1. Department of Informatics, University of Oslo, Norway; 2. Microsoft Norway, Norway; 3. Sleep Laboratory, Lovisenberg Diaconal Hospital, Norway.

Abstract: Pediatric sleep-apnea screening may benefit from AI, but clinicians need to know whether predictions and explanations are reliable. We developed TRUST-Sleep, which analyzes respiratory effort, airflow, pulse rate, and oxygen saturation signals and combines several deep-learning models to provide prediction support for clinical review. It assesses prediction support, explanation reliability, and the need for clinical review using three trust measures: Clinical Decision Trust (CDT), Clinical Explanation Reliability Aggregate (CERA), and Clinical Acceptance Score (CAS). When tested on a pediatric dataset of more than 240,000 sleep-study windows, the system was effective at distinguishing apnea-related from non-apnea windows. It achieved high overall discrimination, while also showing that some event types and the imbalance between common and rare classes remain challenging. Importantly, the trust measures provided a structured way to identify how well supported each prediction and explanation was, rather than presenting every model output as equally reliable. We also tested TRUST-Sleep on an adult sleep dataset to examine how it behaves when applied to substantially different patients, recording conditions, and clinical labels. As expected, prediction performance decreased under this major dataset shift, especially for distinguishing apnea subtypes. However, the uncertainty estimates remained informative, suggesting that the framework can still signal when its predictions are less certain outside the population on which it was developed. At this stage, TRUST-Sleep supports clinician review rather than replacing clinical decisions. It provides calibrated predictions, transparent explanations, and trust indicators to show when outputs are well supported or need caution. Autonomous acceptance of high-trust cases is an ongoing research direction and requires further pediatric validation and workflow studies.