

Nordic AI Meet 2026

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

2. From population norms to personalized trajectories: interpretable Bayesian forecasting for cognitive decline

Authors: Maria Sahakyan¹ Brita Elvevåg^{1,2}

Affiliations: 1. Department of Clinical Medicine, UiT-The Arctic University of Norway, Tromsø, Norway 2. The Norwegian Center for Clinical Artificial Intelligence, University Hospital of North Norway

Abstract: Cognitive decline varies across individuals in both timing and rate, yet it is often interpreted using fixed population thresholds. These thresholds show how a person compares with a reference group, but not whether a new score represents meaningful change from that person's expected trajectory. Meaningful deterioration may therefore remain undetected, particularly in individuals who begin above average and remain within the normative range as they decline. Longitudinal models capture within-person change, but reliable person-specific estimates require repeated observations over time. At the start of monitoring, however, only one or two cognitive assessments may be available. This limited personal history creates a cold-start problem, familiar from recommender systems. To address it, we introduce Personalized Risk Inference via Sequential Monitoring (PRISM), an interpretable framework that combines a population-informed starting point with sequential person-specific updating. An Explainable Boosting Machine (EBM) estimates each person's expected baseline performance on the memory assessment from routine demographic and health variables. As repeated memory scores accumulate, Bayesian updating gives increasing weight to the individual's own data, while temporal decay gives more weight to recent scores. This EBM prediction provides the initial reference. After the first memory score is observed, posterior mean establishes a personal anchor adjusted over time for expected age-related change. Posterior distributions quantify uncertainty in cognitive level and the probability of decline. We evaluated PRISM in the Health and Retirement Study (HRS; N=30,664) and validated it in the Alzheimer's Disease Neuroimaging Initiative (ADNI; N=1,866). In held-out HRS data, PRISM produced modest reductions in one-step-ahead error when forecasting the next observed memory score, relative to demographic-norm and cumulative-average baselines. By contrast, mean per-visit AUC for separating worsening from stable trajectories was 0.85, compared with 0.69 for a standard linear mixed-effects model. The gap was widest early in follow-up, when sparse histories led to stronger shrinkage toward the population trajectory. Among sustained decliners, PRISM flagged 31% of HRS and 41% of ADNI participants at least one visit before their first observed worsening, with median lead times of 6 and 2 years. By the transition visit, these proportions reached 68% and 56%. This contrast suggests that forecasting the next score and detecting meaningful change are related but different tasks. Interpretability also played a practical role. The additive EBM's learned age function adjusted expected performance over time without an external normative table. The same approach may be useful in behavioral, industrial, or environmental monitoring, where change is better understood relative to a person's or system's expected trajectory than to a single threshold.