Leveraging Structural Priors and Historical Data for Practical Treatment Personalization with Multi-Armed Bandits

Newton Mwai

The seminar will be held in

Lecture Room EC, EDIT Building
Hörsalsvägen 11,
Chalmers University of Technology, Campus Johanneberg

on

October 10th, 2025, 13:00

Opponent: Sandeep Juneja, Professor, Ashoka University, India
Grading Committee: Yevgeny Seldin, Professor, University of Copenhagen, **Branislav Kveton**, Principal Research Scientist, Adobe Research, **Slawomir Nowaczyk**, Professor, Halmstad University
Stand-in: Moa Johansson, Docent, Chalmers **Chair: Devdatt Dubhashi**, Professor, Chalmers

The thesis is available at:
Department of Computer Science and Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2025

Phone: +46 (0)31 772 1000

CHALMERS
UNIVERSITY OF TECHNOLOGY

Leveraging Structural Priors and Historical Data for Practical Treatment Personalization with Multi-Armed Bandits

Newton Mwai

Department of Computer Science and Engineering
Chalmers University of Technology

Abstract

Personalizing treatments for patients often requires sequentially trying different options from a set of available therapies until the most effective one is identified for the patient's characteristics. In chronic diseases such as Alzheimer's Disease, where interventions mainly have short-term effects, this search process can be formulated as a multi-armed bandit (MAB) problem. Reducing the length of the search is essential to limit patient burden and other associated costs, while practical constraints, such as limiting switches between therapies, introduce additional complexity to exploration. This thesis advances the foundational understanding and applications of MAB algorithms in the context of treatment personalization, focusing on improving sample efficiency by leveraging latent structure revealed from historical data, and accommodating practical treatment switching constraints. Key contributions include: (i) latent bandit algorithms for fixed-confidence pure exploration, providing new insights into exploration dynamics; (ii) the Identifiable Latent Bandit framework, which learns reward models from observational data under identifiability assumptions; and (iii) Latent Preference Bandits, which relax structural requirements by modeling preference orderings instead of full reward vectors. The work addresses the challenge of switching constraints through batched exploration approaches. Furthermore, the Alzheimer's Disease Causal estimation Benchmark (ADCB), a semi-synthetic simulator integrating real-world Alzheimer's data with domain expertise is designed and employed as a causally sound evaluation platform for bandit algorithms in personalized medicine. Together, these contributions connect theoretical MAB developments with clinically motivated constraints, offering methodologies for more efficient and practical treatment personalization.

Keywords

Treatment personalization, multi-armed bandits, fixed-confidence pure exploration, latent bandits, structural priors, policy learning with historical data, exploration with switching constraints, healthcare bandit simulators