Explainable Health Data Analytics with Tsetlin Machines

Principal supervisor: Dr. Ekin Can Erkus, Electrical and Electronics Engineering, Newcastle University

Email: ekin.erkus@newcastle.ac.uk

Co-supervisor: Dr. Alex Chan, Electrical and Electronics Engineering, Newcastle University

Email: alex.chan@newcastle.ac.uk

Executive Summary

This project builds a reproducible, HPC-enabled pipeline for explainable health data analytics using routinely collected hospital records, with MIMIC-IV as a public testbed. The intern will train Tsetlin Machines to learn transparent rule-based predictors and will benchmark them against standard baselines under matched preprocessing. The evaluation includes fairness across patient subgroups, and explanation stability under missingness and resampling. Outputs include an auditable codebase, automated explainable model that supports deployable, trustworthy prediction from healthcare data.

Project Description

Context and overarching goals

Routinely collected hospital data can support risk stratification and treatment planning, but many predictive models remain difficult to audit and unreliable under missingness and subgroup variation. This project develops an end-to-end, reproducible framework for explainable health data analytics, using MIMIC-IV as a public testbed. The goal is to produce calibrated predictions together with rule-based explanations that can be inspected and stress-tested, and to report fairness-relevant performance across patient subgroups. The scope emphasizes reusable tooling and workflow design rather than a single narrow prediction task.

A leakage-safe pipeline will be implemented for cohorting, label construction, temporal windowing when needed, and feature generation from routinely recorded variables such as demographics, diagnoses, procedures, laboratory values, medications, and bedside observations. Missingness will be handled explicitly using indicator and recency features, and all data-dependent transforms will be fit on training or validation data only. Tsetlin Machines will be the primary model class, producing propositional clause sets that provide both global summaries and instance-level rationales. Performance and calibration will be benchmarked against standard baselines, including logistic regression and gradient-boosted trees, under matched preprocessing and split logic. Explainability will be evaluated quantitatively using rule sparsity and stability under bootstrap resampling and controlled missingness injection. Fairness analysis will report subgroup calibration and error disparities across available demographic strata.

Computational aspects and why HPC resources are needed

The workflow requires repeated large-scale feature construction and model training across multiple feature sets, discretization choices for rule learning, and resampling-based stability evaluations. HPC resources enable parallel hyperparameter search, systematic missingness stress tests, and bootstrap uncertainty estimation within the internship timeframe. HPC workflows will be implemented with containerized execution, configuration tracking, and automated result aggregation, with weekly input from both supervisors.

Planned outputs by the end of the internship

An auditable, version-controlled, HPC-ready codebase for data processing, leakage-safe splitting, training, and evaluation. An automated explainable modeling workflow that outputs rule-based explanations, stability summaries under missingness and resampling, and subgroup performance reports. A concise benchmark report comparing Tsetlin Machines to standard baselines under matched protocols, with guidance for deployable, trustworthy prediction from routinely collected healthcare data.