MACHINE LEARNING FOR HEALTH

Problem Formulation and the Model Life Cycle

How to turn a clinical need into a valid ML study

TECHNICAL HANDOUT · SESSION 04

Professor Flávio Luiz Seixas
Graduate Program in Computing

Academic teaching material based exclusively on the PDFs available in the project

Table of Contents

Introduction

The validity of a machine-learning study begins before the algorithm. A clinical need must be converted into a verifiable task with a clearly defined population, prediction time, inputs, outcome, horizon, and intended action. [1]

This session organizes these decisions and shows why the cohort, label, temporal windows, and partitioning are part of the system specification rather than a later cleaning step. [1]

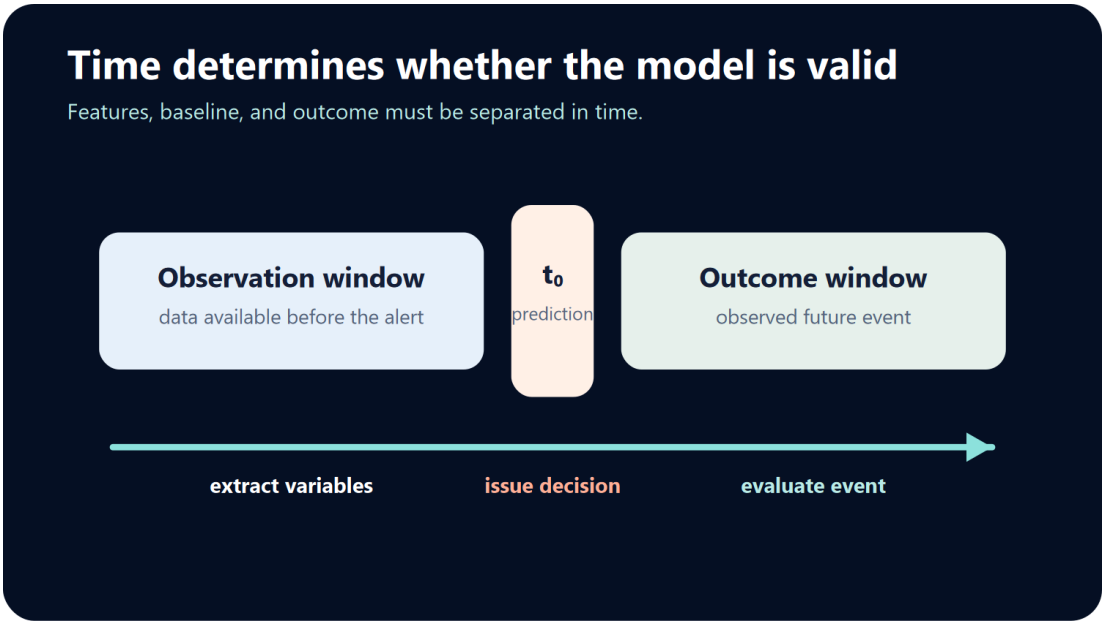

Figure 1 — Observation, prediction, and outcome windows. Source: original illustration based on the course material and [1]–[3].

1. From clinical need to task

An ML task specifies an input and an output. In health care, it should also indicate who will receive the result, when it will be available, and which decision it may support. Different problems—diagnosis, prognosis, detection, and prioritization—require different targets, horizons, and evaluations [1].

The cohort delimits the population in which the model is developed. Inclusion and exclusion criteria are operational hypotheses about who represents the population of use. Electronic phenotyping combines codes, measurements, procedures, text, or temporality to identify patients and can therefore also introduce classification error [1].

KEY CONCEPT
An ML task specifies an input and an output.

2. Labels and reference standards

The label is the operational representation of the desired outcome. It may come from adjudication, a test, a code, or a clinical action; each source has its own error and bias. When no gold standard exists, uncertainty in the reference must be documented and, whenever possible, evaluated with curated

samples [2].

Labels extracted from care may contain prior decisions. A treatment, confirmatory test, or referral may reveal the outcome while also being a consequence of the process the model is intended to anticipate [1].

3. Temporal windows and partitioning

The observation window contains only information available up to the prediction time; the outcome window begins afterward. Separating baseline, observation, and horizon prevents the future from crossing the decision boundary. The same patient should not appear on different sides of a split when this would permit memorization of individual characteristics [1].

Training, validation, and test sets serve different purposes. Splits by time and institution expose changes that a local random split may conceal. The test set must remain independent of the choices made during development [2].

4. From protocol to monitoring

A minimum protocol describes the question, cohort, index time, predictors, outcome, partitioning, metrics, subgroups, and analysis plan. After deployment, the population and processes continue to change; the cycle therefore includes monitoring and review and does not end at deployment [3].

Session summary

From clinical need to task: an ml task specifies an input and an output.
Labels and reference standards: the label is the operational representation of the desired outcome.
Temporal windows and partitioning: the observation window contains only information available up to the prediction time; the outcome window begins afterward.
From protocol to monitoring: a minimum protocol describes the question, cohort, index time, predictors, outcome, partitioning, metrics, subgroups, and analysis plan.

References

[1] Subramanian, D.; Cohen, T. A. Machine Learning Systems. In: Intelligent Systems in Medicine and Health. Springer, 2022. External link

Local file: pdfs/3-Intelligent Systems/VI-Machine Learning Systems.pdf · PDF pages used: 3, 4, 10, 12, 16, 67.

[2] Rajkomar, A.; Dean, J.; Kohane, I. Machine Learning in Medicine. New England Journal of Medicine, 2019.

Local file: pdfs/1-Machine Learning/NEJMra1814259.pdf · PDF pages used: 2, 3, 8, 9.

[3] Bellazzi, R.; Dagliati, A.; Nicora, G. Predicting Medical Outcomes. In: Intelligent Systems in Medicine and Health. Springer, 2022. External link

Local file: pdfs/3-Intelligent Systems/XI-Predicting Medical Outcomes.pdf · PDF pages used: 3, 4, 21, 23, 24.