

APRENDIZADO DE MÁQUINA PARA SAÚDE

Formulação do Problema e Ciclo de Vida do Modelo

Como transformar uma necessidade clínica em um estudo de ML válido

APOSTILA TÉCNICA · ENCONTRO 04

Prof. Flávio Luiz Seixas
Pós-graduação em Computação

Sumário

Sumário	2
Introdução	3
1. Da necessidade clínica à tarefa	3
2. Rótulos e referência	3
3. Janelas temporais e particionamento	4
4. Do protocolo ao monitoramento.....	4
Síntese do encontro.....	4
Referências	4

Introdução

A validade de um estudo de aprendizado de máquina começa antes do algoritmo. Uma necessidade clínica precisa ser convertida em tarefa verificável, com população, instante de previsão, entradas, desfecho, horizonte e ação pretendida claramente definidos. [1]

Este encontro organiza essas decisões e mostra por que coorte, rótulo, janelas temporais e particionamento fazem parte da especificação do sistema, não de uma limpeza posterior. [1]

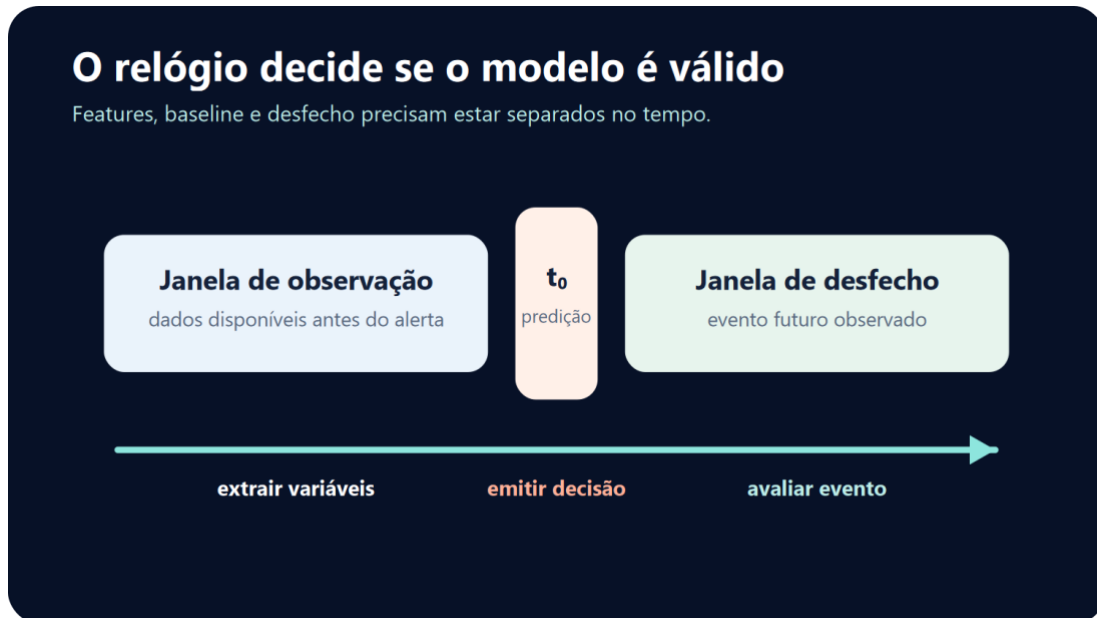


Figura 1 — Janelas de observação, previsão e desfecho. Fonte: elaboração própria a partir do material da disciplina, com base em [1]–[3].

1. Da necessidade clínica à tarefa

Uma tarefa de ML especifica uma entrada e uma saída. Em saúde, também deve indicar quem receberá o resultado, quando ele estará disponível e que decisão poderá apoiar. Problemas diferentes — diagnóstico, prognóstico, detecção e priorização — requerem alvos, horizontes e avaliações diferentes [1].

A coorte delimita o universo no qual o modelo é desenvolvido. Critérios de inclusão e exclusão são hipóteses operacionais sobre quem representa a população de uso. Fenotipagem eletrônica combina códigos, medidas, procedimentos, texto ou temporalidade para identificar pacientes, e por isso também pode introduzir erro de classificação [1].

CONCEITO-CHAVE

Uma tarefa de ML especifica uma entrada e uma saída.

2. Rótulos e referência

O rótulo é a representação operacional do resultado desejado. Ele pode vir de adjudicação, teste, código ou ação clínica; cada origem tem erro e viés próprios. Quando não há padrão-ouro, a incerteza da referência precisa ser documentada e, quando possível, avaliada com amostras curadas [2].

Rótulos extraídos do cuidado podem conter decisões anteriores. Um tratamento, exame confirmatório ou encaminhamento pode revelar o desfecho e, simultaneamente, ser consequência do processo que se pretende antecipar [1].

3. Janelas temporais e particionamento

A janela de observação contém somente informações disponíveis até o instante de previsão; a janela de desfecho começa depois dele. Separar baseline, observação e horizonte impede que o futuro atravesse a fronteira da decisão. O mesmo paciente não deve aparecer em lados diferentes da divisão quando isso permitir memorização de características individuais [1].

Conjuntos de treino, validação e teste cumprem papéis distintos. Separações por tempo e instituição expõem mudanças que uma divisão aleatória local pode esconder. O teste deve permanecer independente das escolhas feitas durante o desenvolvimento [2].

4. Do protocolo ao monitoramento

Um protocolo mínimo descreve pergunta, coorte, índice temporal, preditores, desfecho, particionamento, métricas, subgrupos e plano de análise. Depois da implantação, a população e os processos continuam mudando; por isso, o ciclo inclui monitoramento e revisão, e não termina no deploy [3].

Síntese do encontro

Da necessidade clínica à tarefa: uma tarefa de ml especifica uma entrada e uma saída.
Rótulos e referência: o rótulo é a representação operacional do resultado desejado.
Janelas temporais e particionamento: a janela de observação contém somente informações disponíveis até o instante de previsão; a janela de desfecho começa depois dele.
Do protocolo ao monitoramento: um protocolo mínimo descreve pergunta, coorte, índice temporal, preditores, desfecho, particionamento, métricas, subgrupos e plano de análise.

Referências

- [1] Subramanian, D.; Cohen, T. A. Machine Learning Systems. In: Intelligent Systems in Medicine and Health. Springer, 2022.
- [2] Rajkomar, A.; Dean, J.; Kohane, I. Machine Learning in Medicine. New England Journal of Medicine, 2019.
- [3] Bellazzi, R.; Dagliati, A.; Nicora, G. Predicting Medical Outcomes. In: Intelligent Systems in Medicine and Health. Springer, 2022.