



CEFET/RJ

Avaliação de Métodos de Detecção



Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

<https://eic.cefet-rj.br/~eogasawara>

Por que Avaliar?

- Avaliar a qualidade dos métodos de detecção é essencial.
- • Permite comparar diferentes abordagens de forma objetiva.
- • Fundamental para aplicações críticas e sensíveis a erro.

Matriz de confusão

- A matriz de confusão organiza os acertos e erros de um método em quatro categorias:
 - Verdadeiros Positivos (VP)
 - Falsos Positivos (FP)
 - Verdadeiros Negativos (VN)
 - Falsos Negativos (FN)
- A base para cálculo de métricas como precisão, revocação e F1

		Detected events		Total
		yes	no	
Known events	yes	TP	FN	P
	no	FP	TN	N
Total		P'	N'	$P + N$

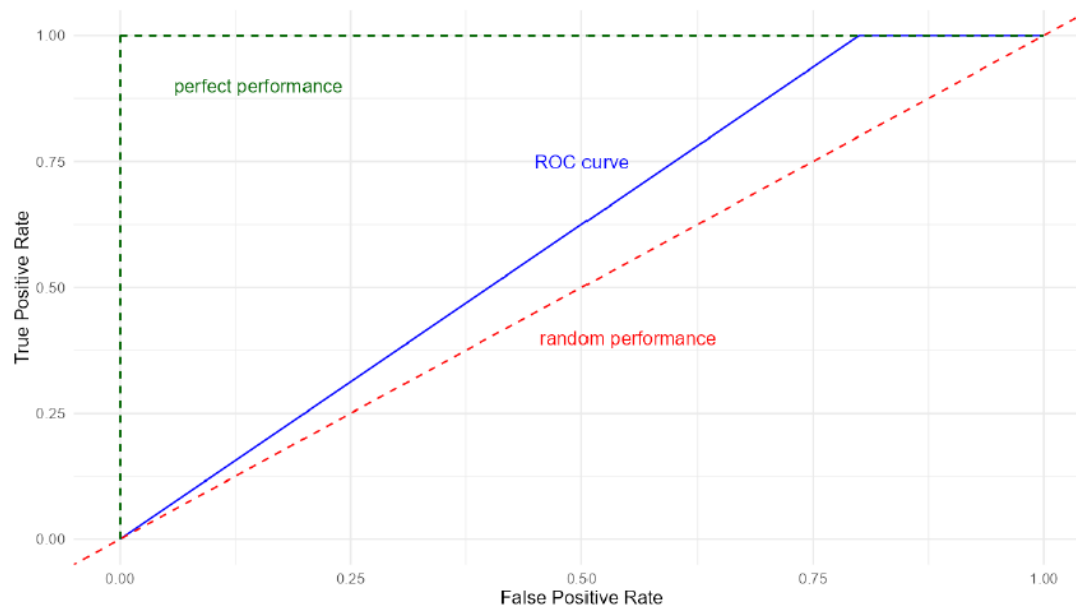
Métricas Clássicas

- Precisão: proporção de verdadeiros positivos entre os detectados.
- Revocação: proporção de verdadeiros positivos entre os eventos reais
- F1-Score: harmônica entre precisão e revocação

Metric	Formula
accuracy, recognition rate	$\frac{TP + TN}{P + N}$
error rate, misclassification rate	$\frac{FP + FN}{P + N}$
precision	$\frac{TP}{TP + FP}$
recall, sensitivity, true positive rate	$\frac{TP}{P}$
F_1 , F-score, harmonic mean of precision and recall	$\frac{2 \times precision \times recall}{precision + recall}$
specificity, true negative rate	$\frac{TN}{N}$
fallout, false positive rate	$\frac{FP}{N}$
F_β , where β is a non-negative real number	$\frac{(1 + \beta^2) \times precision \times recall}{\beta^2 \times precision + recall}$

Curva ROC e AUC

- A curva ROC representa a relação entre taxa de verdadeiros positivos e taxa de falsos positivos em diferentes limiares.
- A área sob a curva (AUC) indica a capacidade de separação do classificador. Quanto maior a AUC, melhor o desempenho.



Tolerância Temporal

- Métodos de pontuação tradicionais, como precisão e revocação, não são suficientes para avaliar o desempenho da detecção de eventos online
- Eles não incorporam o tempo e não recompensam a detecções próximas
 - Verdadeiros positivos são recompensados
- Todos os outros resultados são "severamente" e igualmente punidos

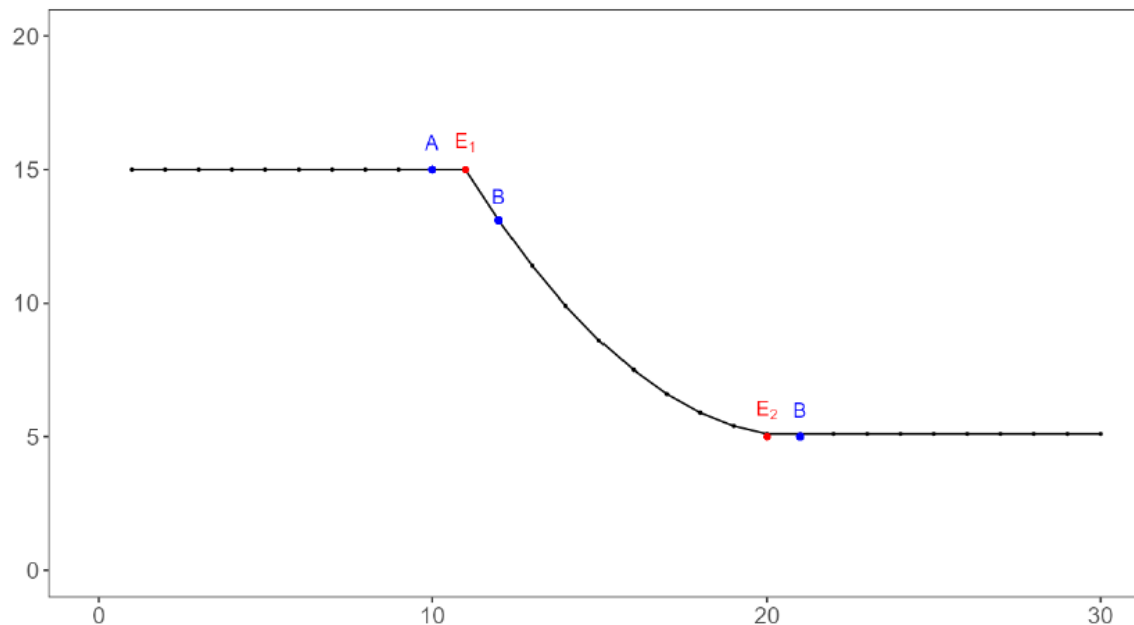


Fig. 6.2 Example regarding the problem of evaluating the detection of events at time 11 (E_1) and 20 (E_2). Method A and B detect an event at times 10 and 12 concerning E_1 . Method B was the only method that detected an event close to E_2 .

Avaliação contínua em fluxo de dados

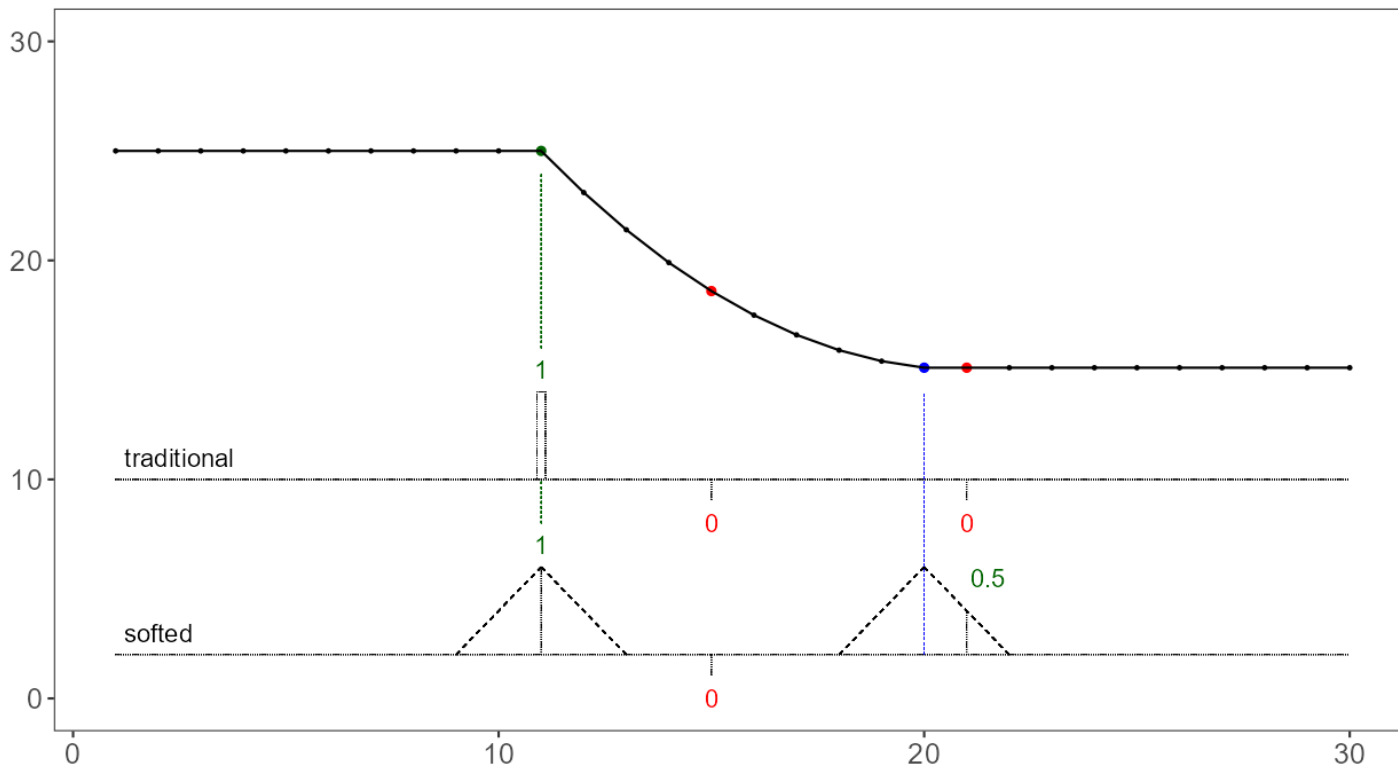
- Em ambientes de detecção online, a avaliação também deve ser feita em tempo real:
 - Monitoramento contínuo da precisão, F1, revocação
 - Avaliação com janelas deslizantes (ex: últimos N eventos)
 - Detecção de queda de desempenho ao longo do tempo (ex: por drift)
 - Essencial para adaptação dinâmica de modelos

Outros Critérios Relevantes

- Latência de detecção
- Falsos positivos e falsos negativos
- Eficiência computacional (tempo e memória)
- Generalização para diferentes domínios e séries

Softed

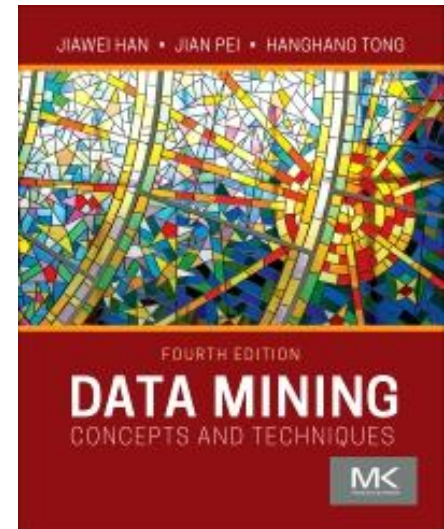
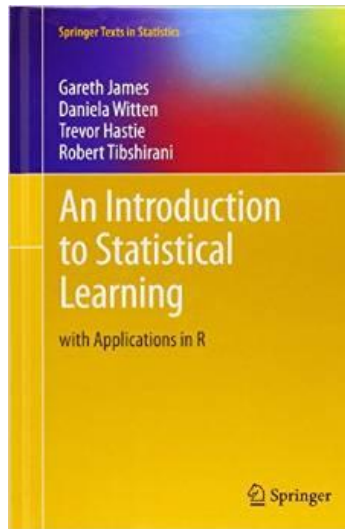
- SoftED introduz uma avaliação mais realista para eventos em séries temporais:
 - Detection Consistency: mede a persistência e regularidade da detecção
 - Detection Mapping: avalia se o evento foi detectado de forma coerente ao longo do tempo
- Penaliza menos as pequenas imprecisões temporais, e mais a inconsistência



Resumo do Capítulo

- Avaliação é fundamental para validar a eficácia dos métodos
- Métricas clássicas (precisão, revocação, F1) são amplamente usadas
- Tolerância temporal torna a avaliação mais realista
- Também devem ser considerados aspectos como latência e eficiência

Referências



- [1] Ogasawara, E.; Salles, R.; Porto, F.; Pacitti, E. Event Detection in Time Series. 1. ed. Cham: Springer Nature Switzerland, 2025.
- [2] Cryer, J. D.; Chan, K.-S. Time Series Analysis: With Applications in R. Springer Science & Business Media, 2008.
- [3] Han, J.; Pei, J.; Tong, H. Data Mining: Concepts and Techniques. 4th edition ed. Cambridge, MA: Morgan Kaufmann, 2022.
- [4] James, G. M.; Witten, D.; Hastie, T.; Tibshirani, R. An Introduction to Statistical Learning: With Applications in R. [s.l.] Springer Nature, 2021.

