



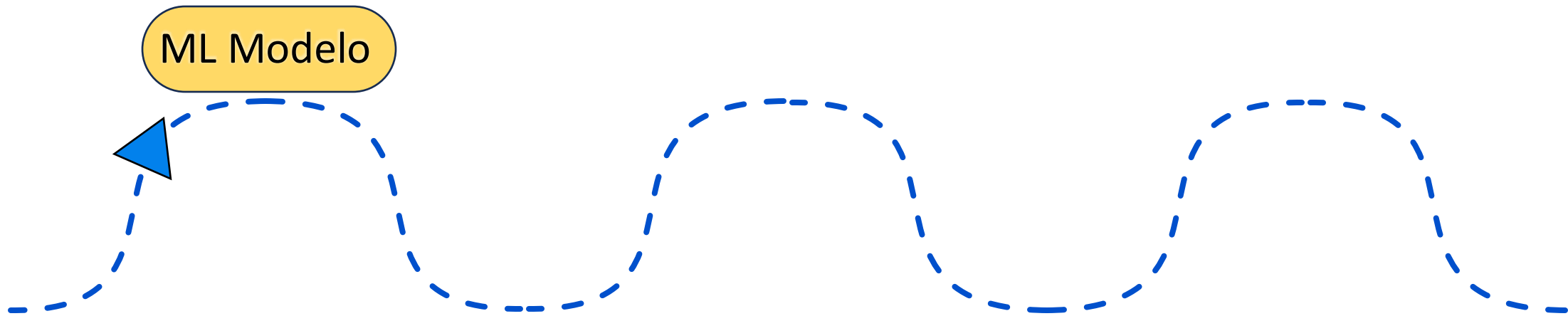
Cornell University

K. Lisa Yang Center for Conservation Bioacoustics

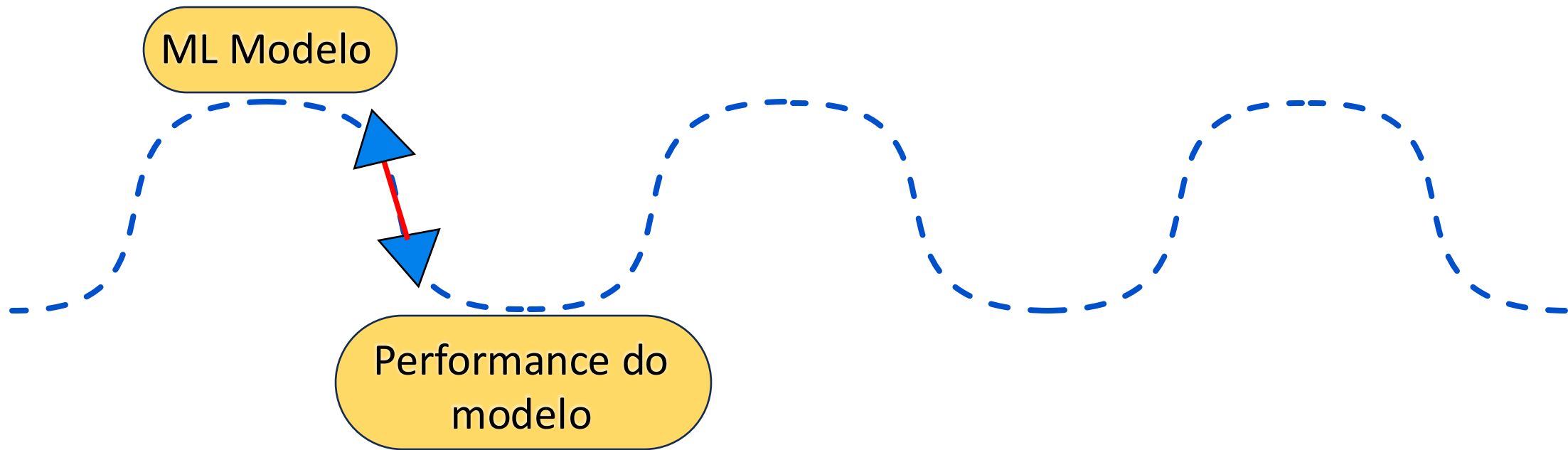
# Ciclo de análises em MAP e performance de modelos

Larissa Sugai

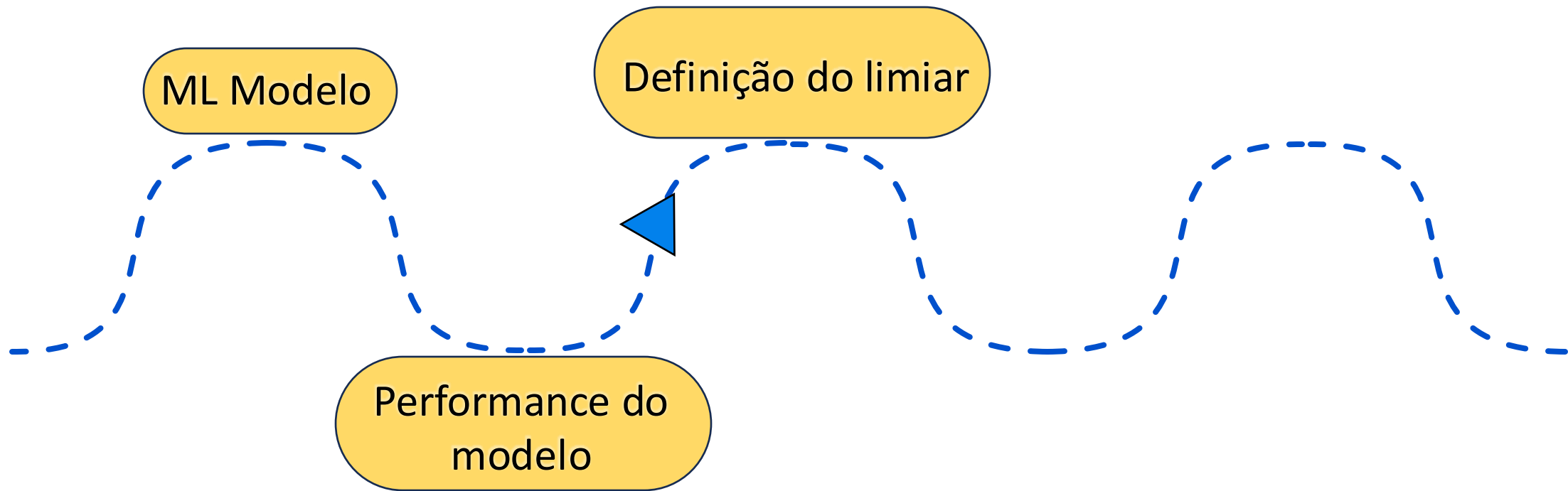
# Workflow para ML em dados bioacústicos



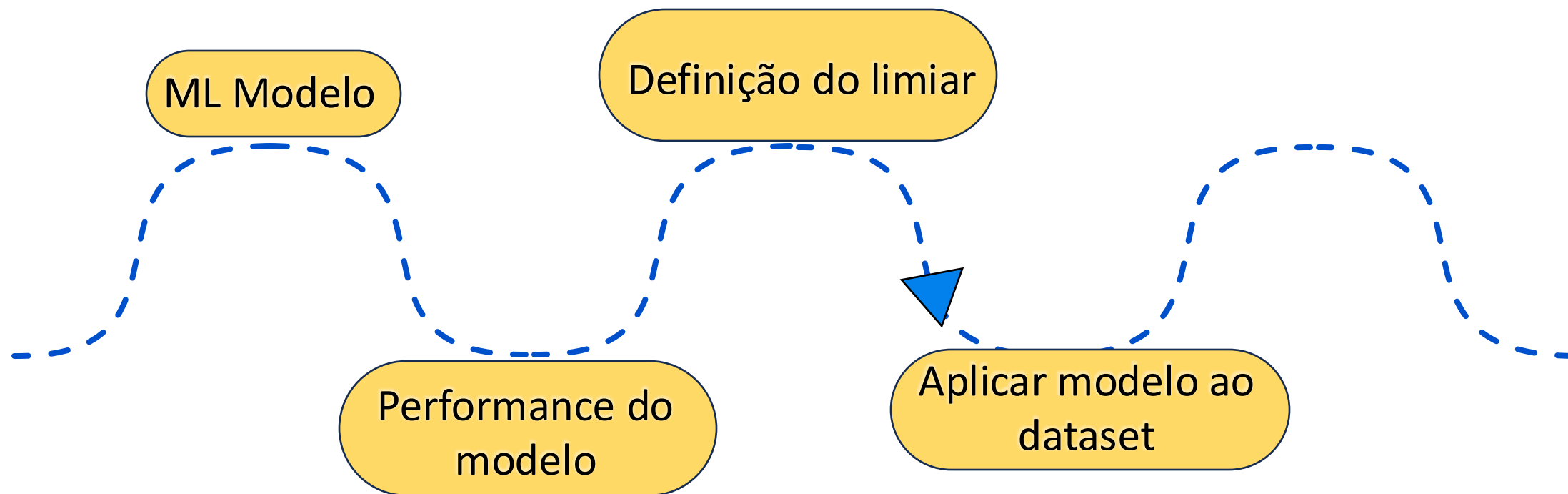
# Workflow para ML em dados bioacústicos



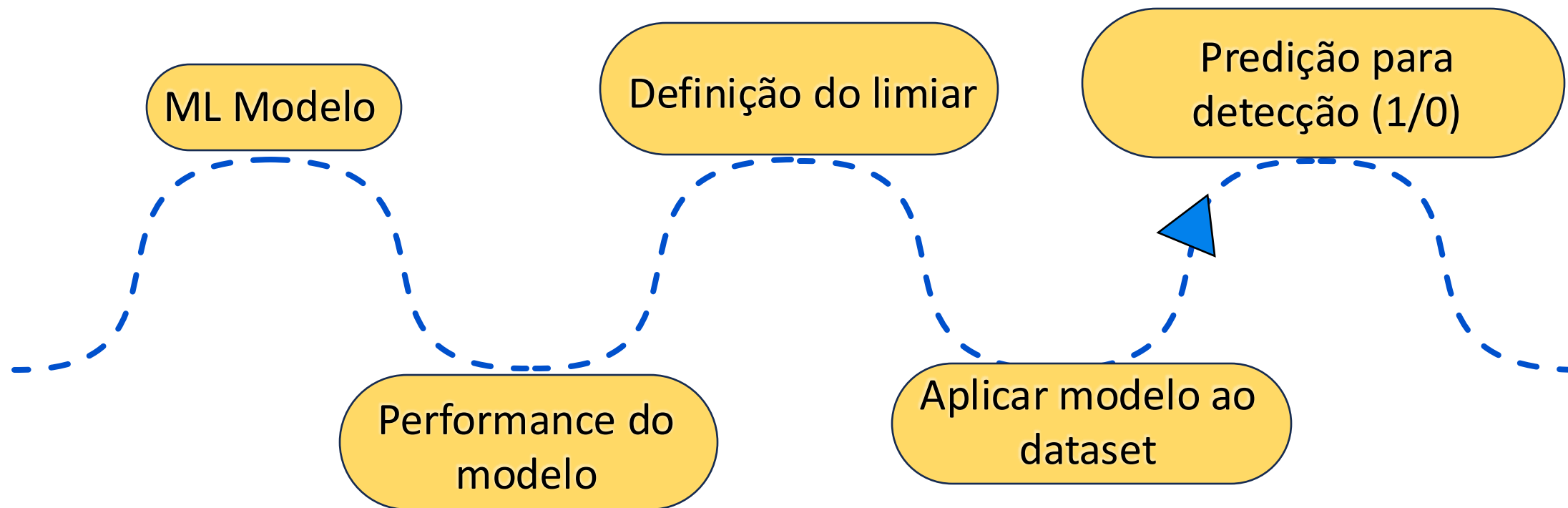
# Workflow para ML em dados bioacústicos



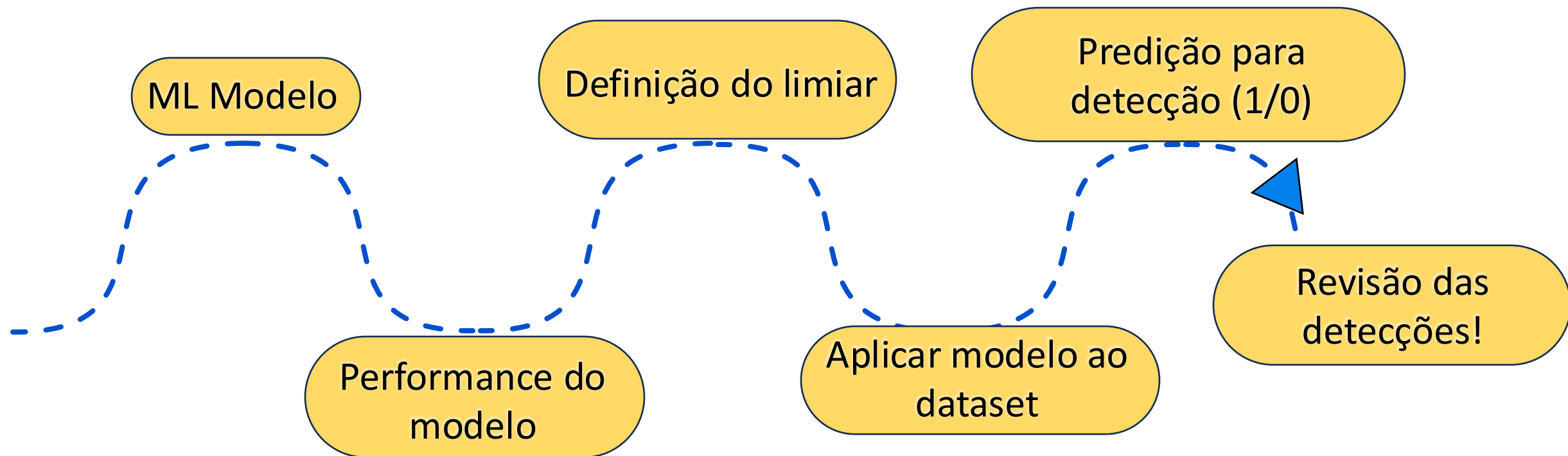
# Workflow para ML em dados bioacústicos



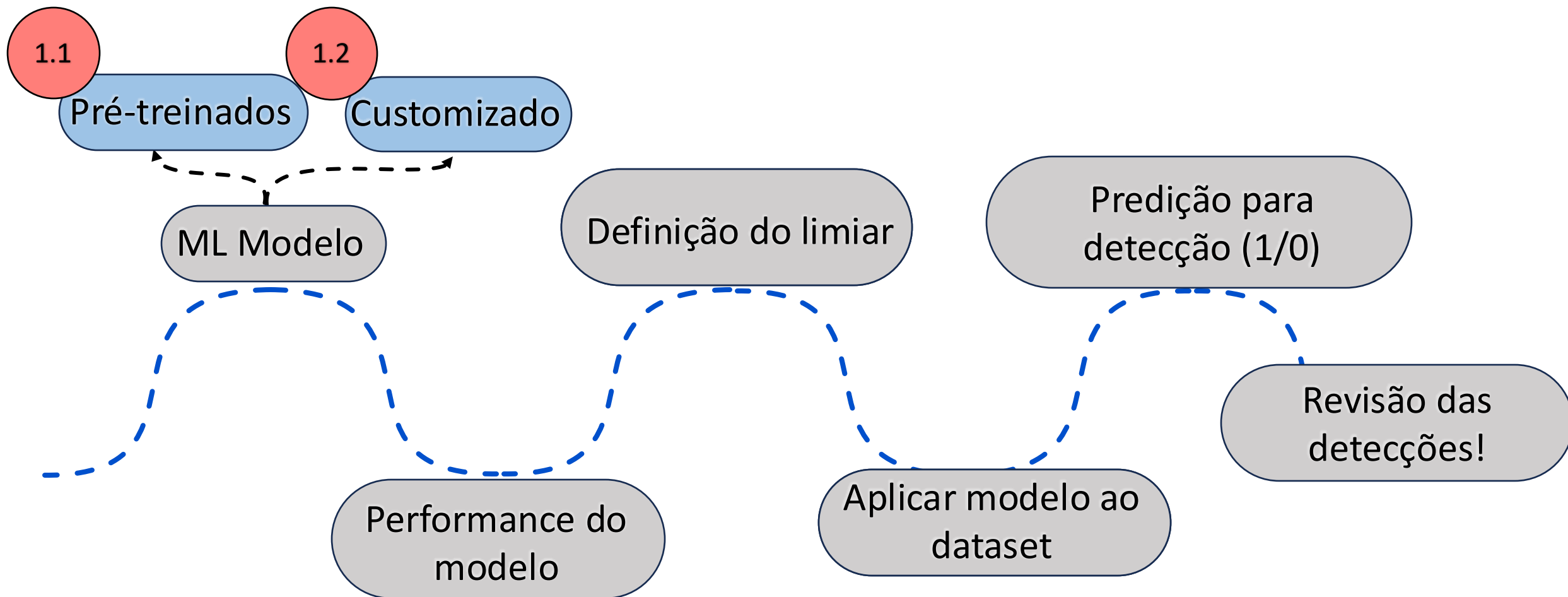
# Workflow para ML em dados bioacústicos



# Workflow para ML em dados bioacústicos



# Workflow para ML em dados bioacústicos

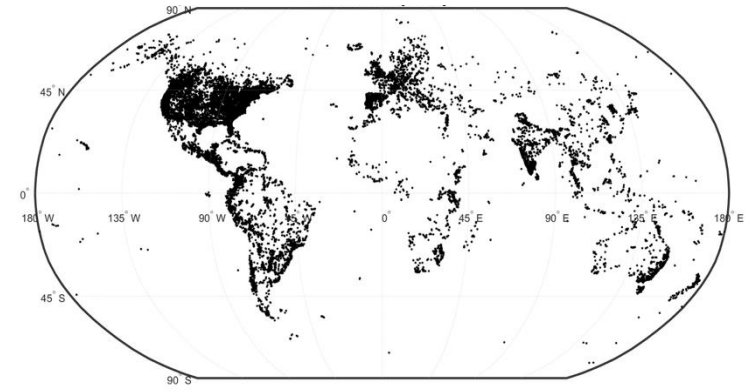




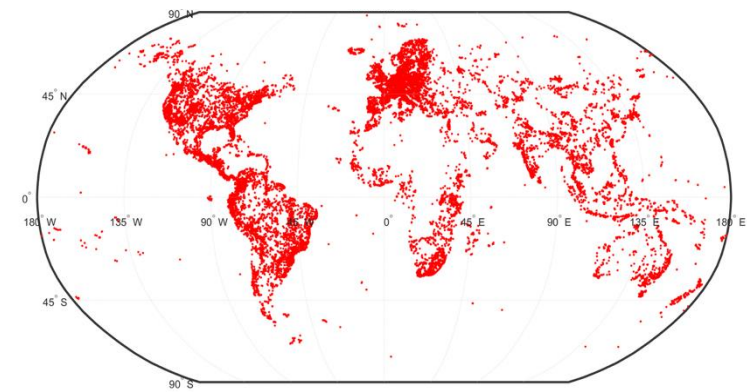
# Modelos pré-treinados

- Contém > 6,000 espécies de ave
- Dados de treinamento: Macaulay Library e Xeno-canto
- MAP > 80%
- Várias plataformas (inclusive o GUI)

~230,000 ML recordings



~400,000 XC recordings



*The Arthur  
Vining Davis  
Foundations*



CHEMNITZ UNIVERSITY  
OF TECHNOLOGY

# Modelos pré-treinados

- Entrada: áudios
- Entrada: lista de espécies ou região
- Output: predições

TheCornellLab

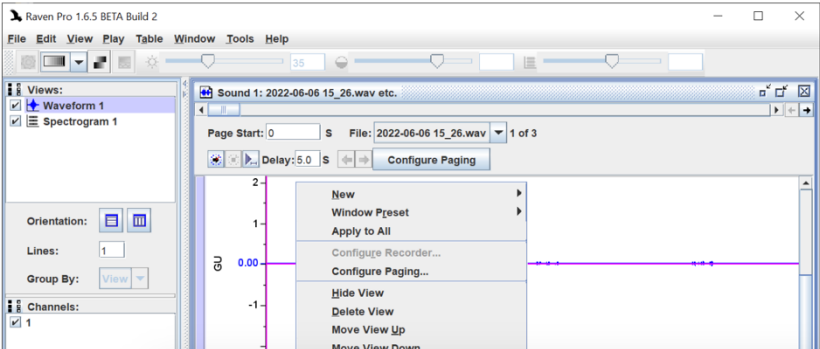
K. Lisa Yang Center for Conservation Bioacoustics

Raven Sound Analysis

Home Software

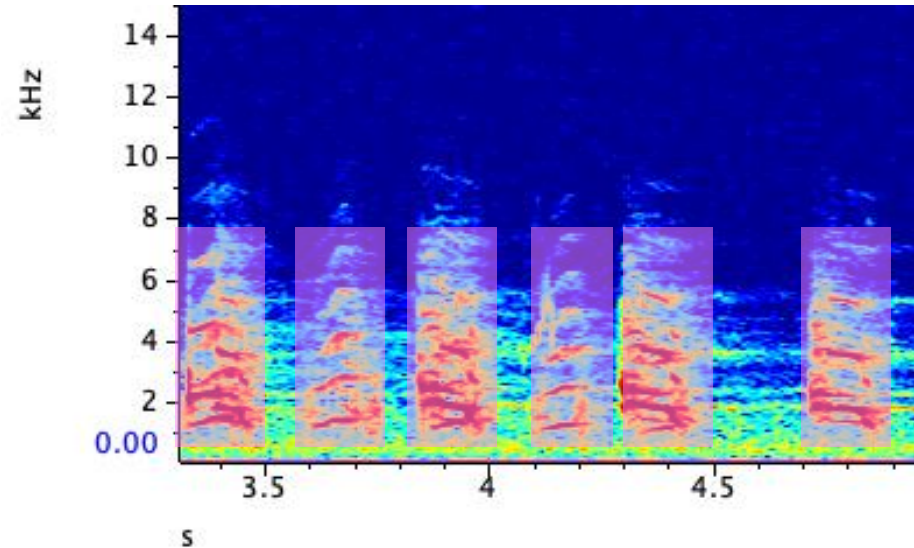
## QUICK START GUIDE

- **Detection Modes:** The learning detector can be invoked in four modes (details in [Chapter 10](#) of the Raven Pro User's Manual):
  - **Interactive**
    - Click on waveform axis label to activate waveform view. Select Learning Detector using View menu.



# A anatomia de uma predição do BirdNET

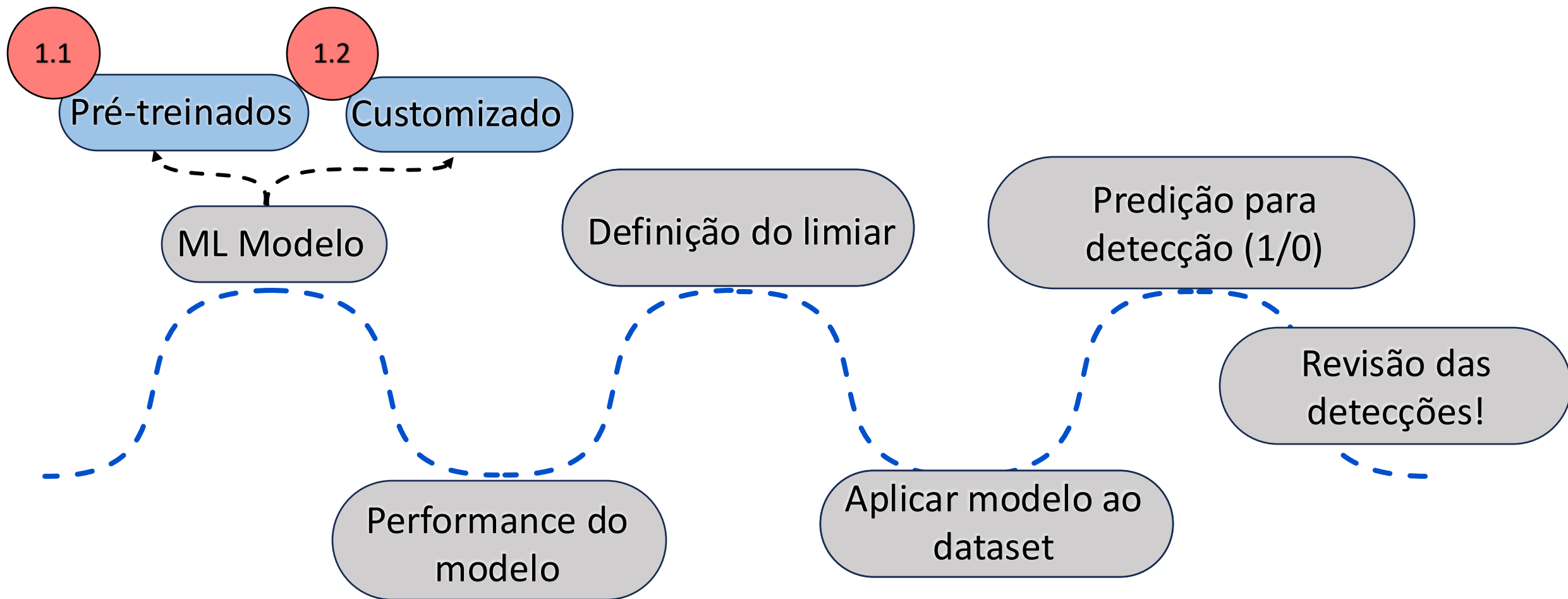
- Escala: 3 segundos
- Frequência: default 0-1500 Hz
- Species Code and Common Name\*
- Confidence score
  - Varia entre 0 - 1
  - **NÃO** é uma probabilidade
  - **NÃO** é comparável entre classes diferentes



Selection Table: S19\_20180214\_080003\_resampled.BirdNET.selection.table.txt

Selection	View	Channel	Begin Time (s)	End Time (s)	Low Freq (Hz)	High Freq (Hz)	Species Code	Common Name	Confidence
<input type="checkbox"/> 1		1	1 111.0000	114.00...	150.000	15000....			0.6280
<input type="checkbox"/> 2		1	1 114.0000	117.00...	150.000	15000....			0.7648
<input type="checkbox"/> 3		1	1 135.0000	138.00...	150.000	15000....			0.5023
<input type="checkbox"/> 4		1	1 183.0000	186.00...	150.000	15000....			0.7398
<input type="checkbox"/> 5		1	1 198.0000	201.00...	150.000	15000....			0.5269
<input type="checkbox"/> 6		1	1 201.0000	204.00...	150.000	15000....			0.5144
<input type="checkbox"/> 7		1	1 1389.00...	1392.0...	150.000	15000....			0.5758
<input type="checkbox"/> 8		1	1 1392.00...	1395.0...	150.000	15000....			0.9818
<input type="checkbox"/> 9		1	1 1395.00...	1398.0...	150.000	15000....			0.9948
<input type="checkbox"/> 10		1	1 1398.00...	1401.0...	150.000	15000....			0.7465

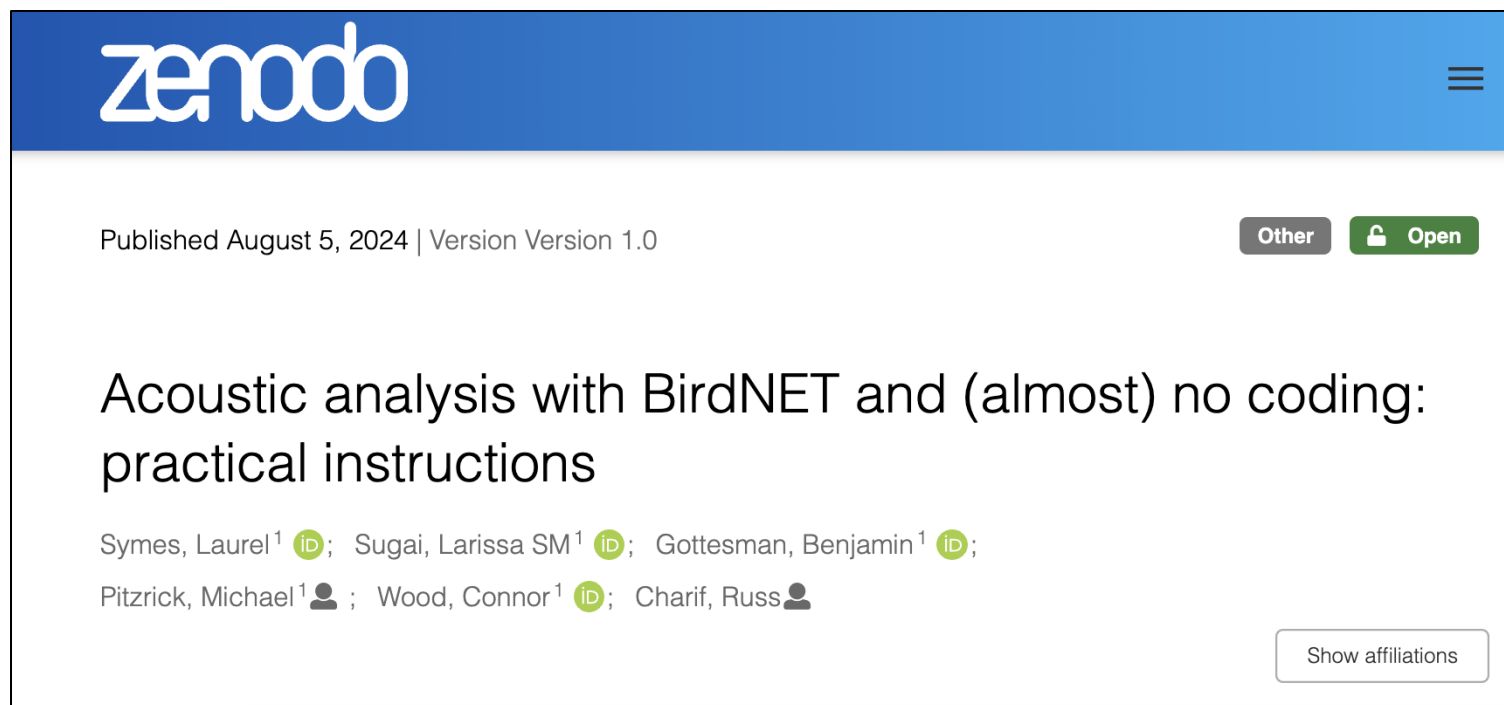
# Workflow para ML em dados bioacústicos



1.2

## Modelos customizados (BirdNET GUI Analyzer)

- Transferência de aprendizado: saindo de um domínio e indo para outro (similar)



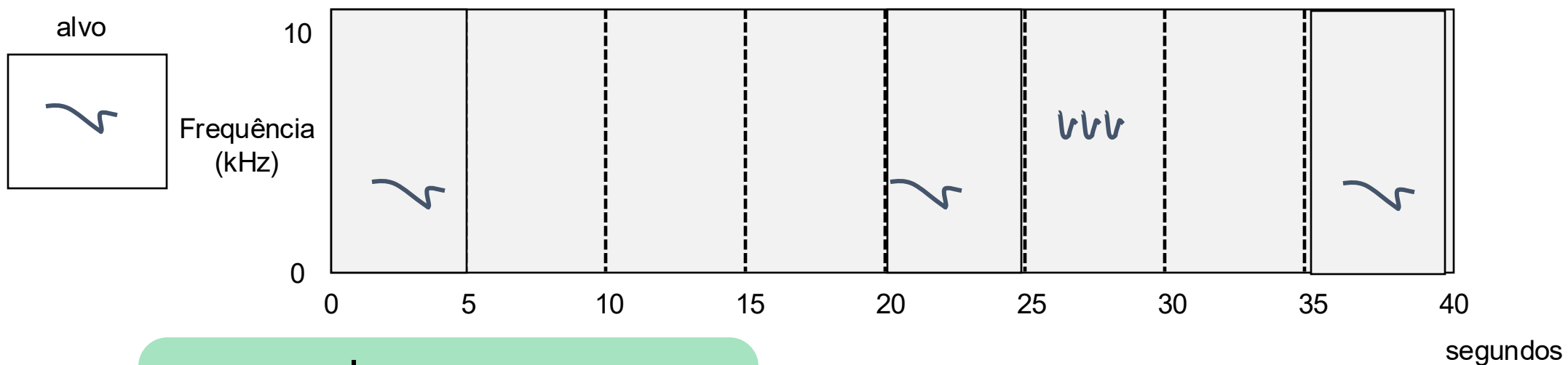
<https://zenodo.org/records/8415090>

## Modelos customizados (BirdNET GUI Analyzer)

- Entrada: dataset de treinamento: clips de 3 segundos
  - Classe alvo (1 ou mais)
    - Dados vindo de gravações manuais / outros datasets
    - Boa qualidade
    - Criado com anotações (caixas de seleção) e clipagem (pode ser feito no Raven)
    - Pode ser menor do que 3 segundos (centraliza e adiciona silêncio)
    - Se possui > 3 segundos, centraliza e corta 3 segundos
    - Atenção com o que está presente nesses 3 segundos
  - Classe 'Other': barulho de fundo, exemplos de falsos positivos
- Controle de qualidade do dataset de treinamento (criar de forma que consiga remover sinais, e.g. anotar qualidade som, tipo de som, etc)!

1.2

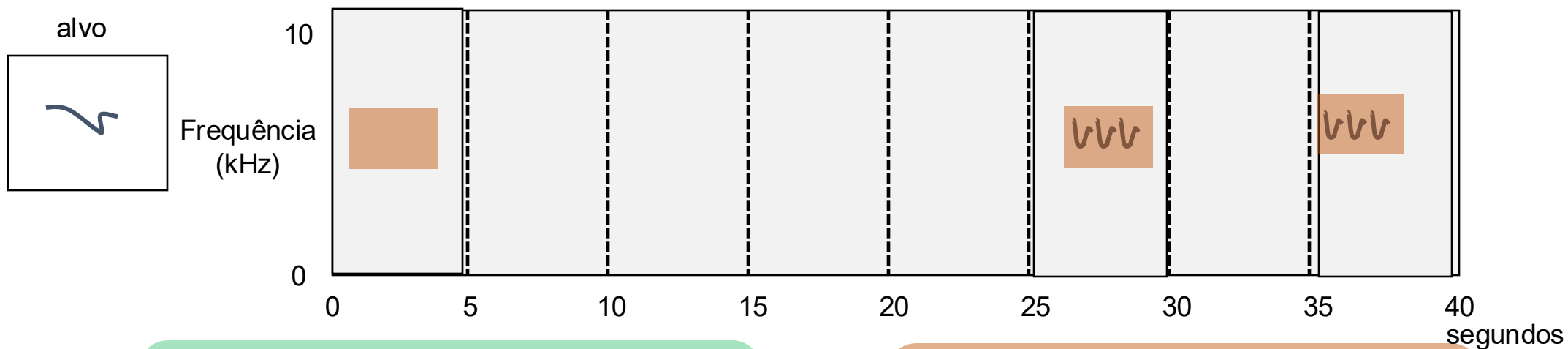
## Modelos customizados (BirdNET GUI Analyzer)



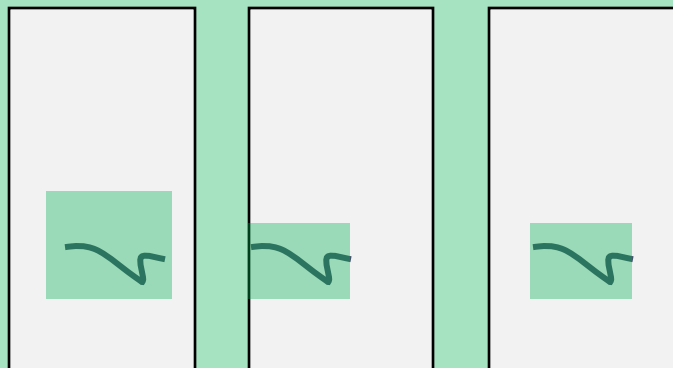
alvo

1.2

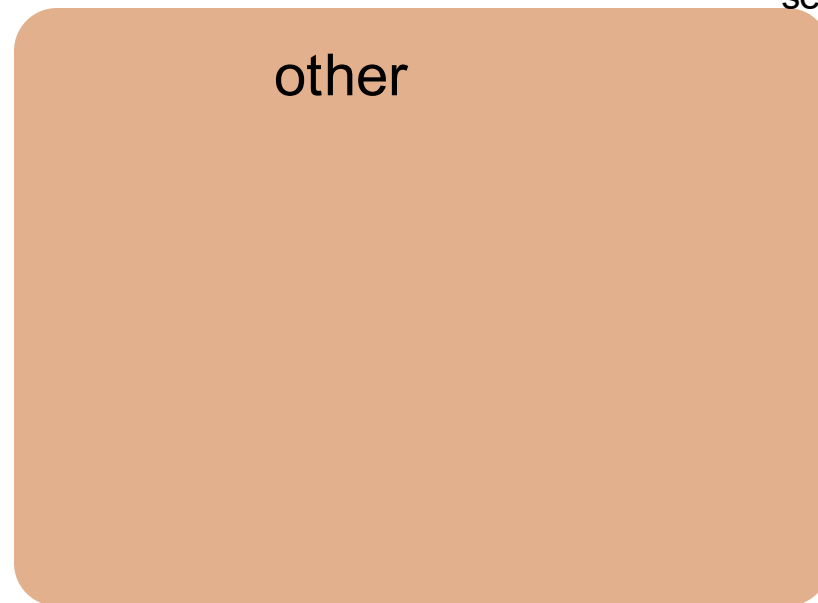
## Modelos customizados (BirdNET GUI Analyzer)



alvo

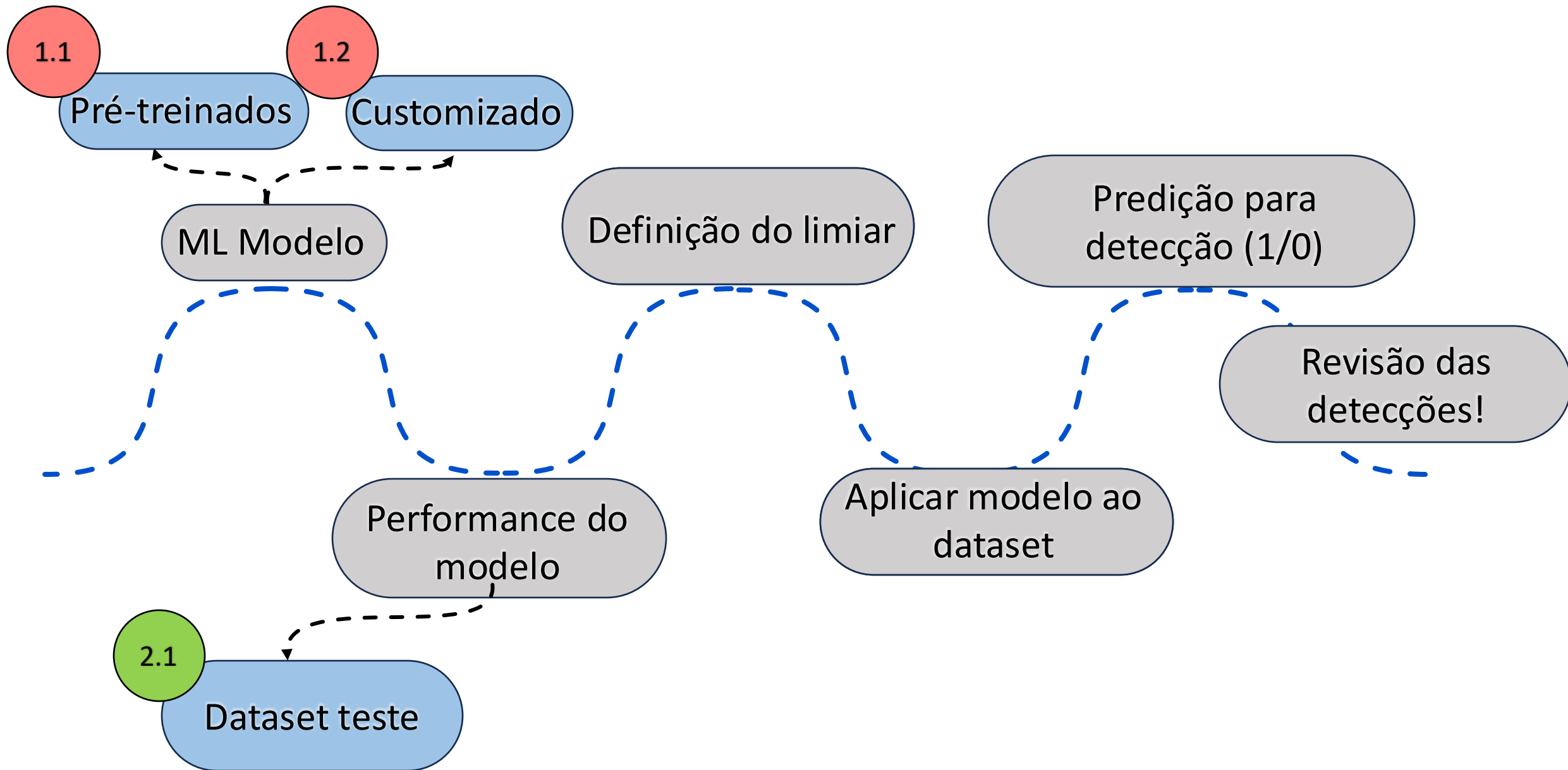


other





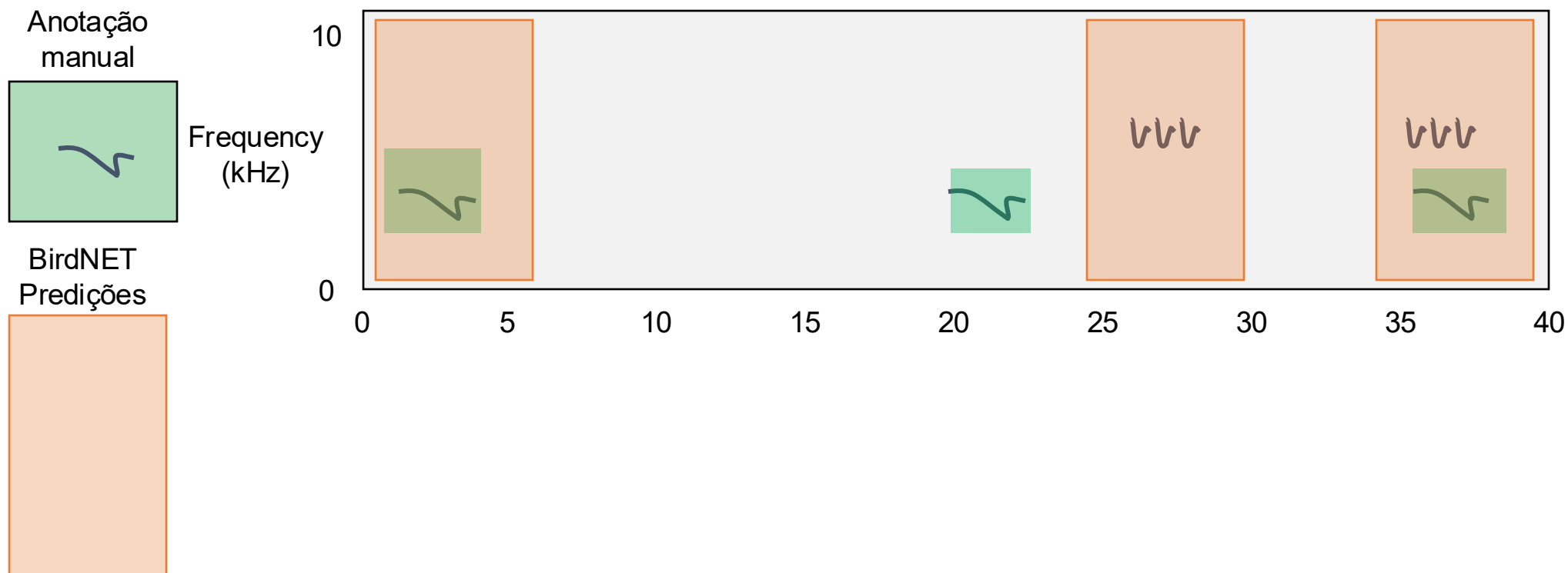
# Workflow para ML em dados bioacústicos



## 2.1

# Avaliação da performance: dataset de teste

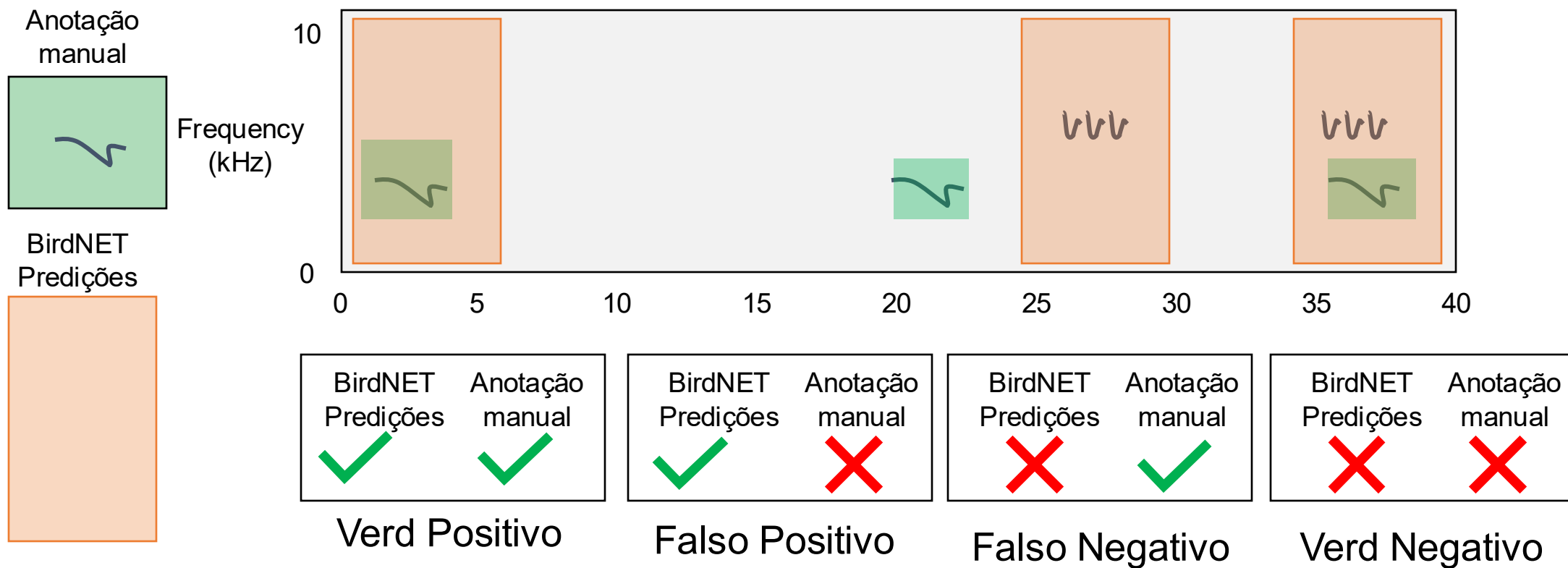
- Um dataset que seja representativo do seu universo amostral (e.g. 5% do total)
- Áudios anotados em total (para todas as classes de interesse)
- Predições do BirdNET são comparadas às anotações manuais



## 2.1

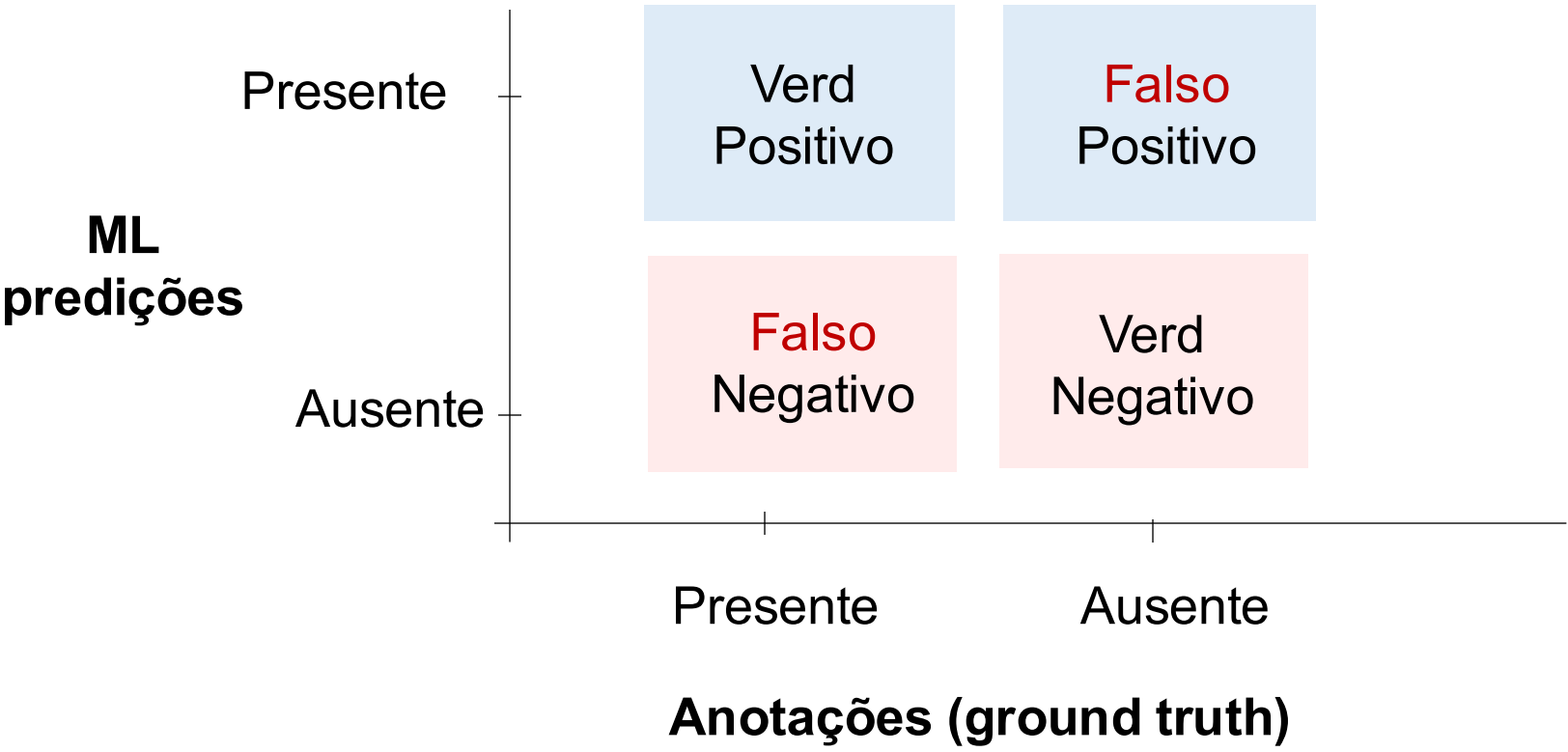
# Avaliação da performance: dataset de teste

- Um dataset que seja representativo do seu universo amostral (e.g. 5% do total)
- Áudios anotados em total (para todas as classes de interesse)
- Predições do BirdNET são comparadas às anotações manuais



2.1

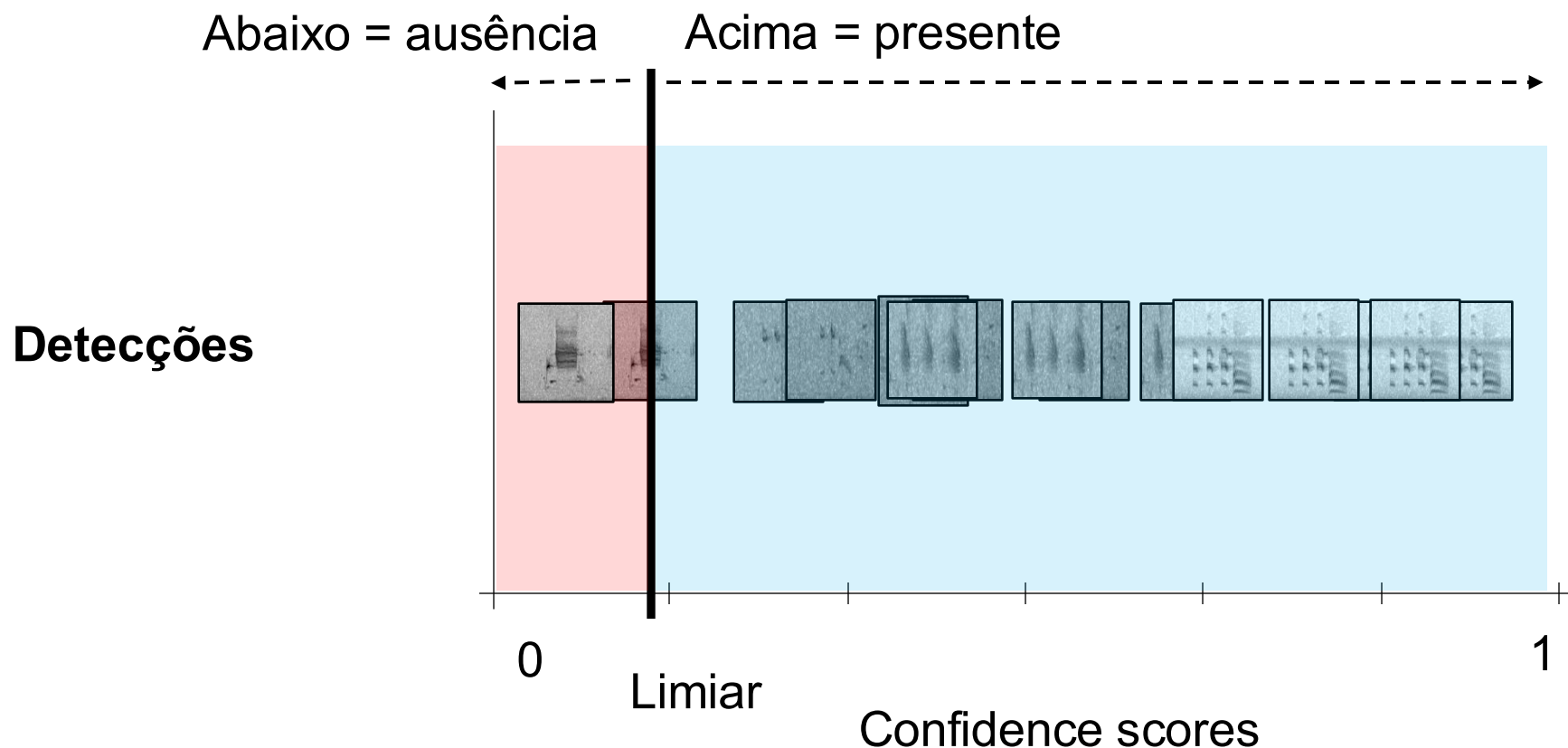
# Avaliação da performance: dataset de teste



## 2.1

## Limiar (threshold): tornando uma predição em detecção

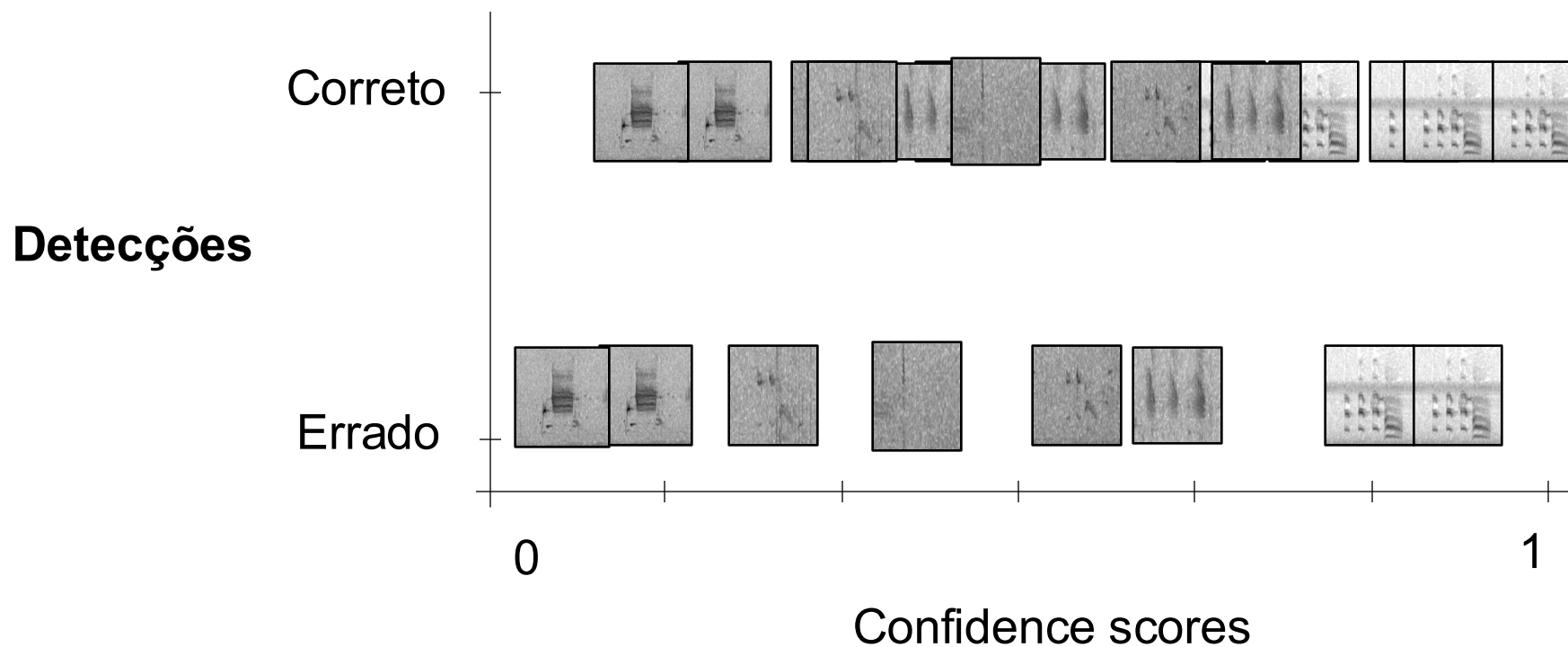
- Critério de corte para classificação: Acima do valor = 1; Abaixo do valor = 0



## 2.1

## Limiar (threshold): tornando uma predição em detecção

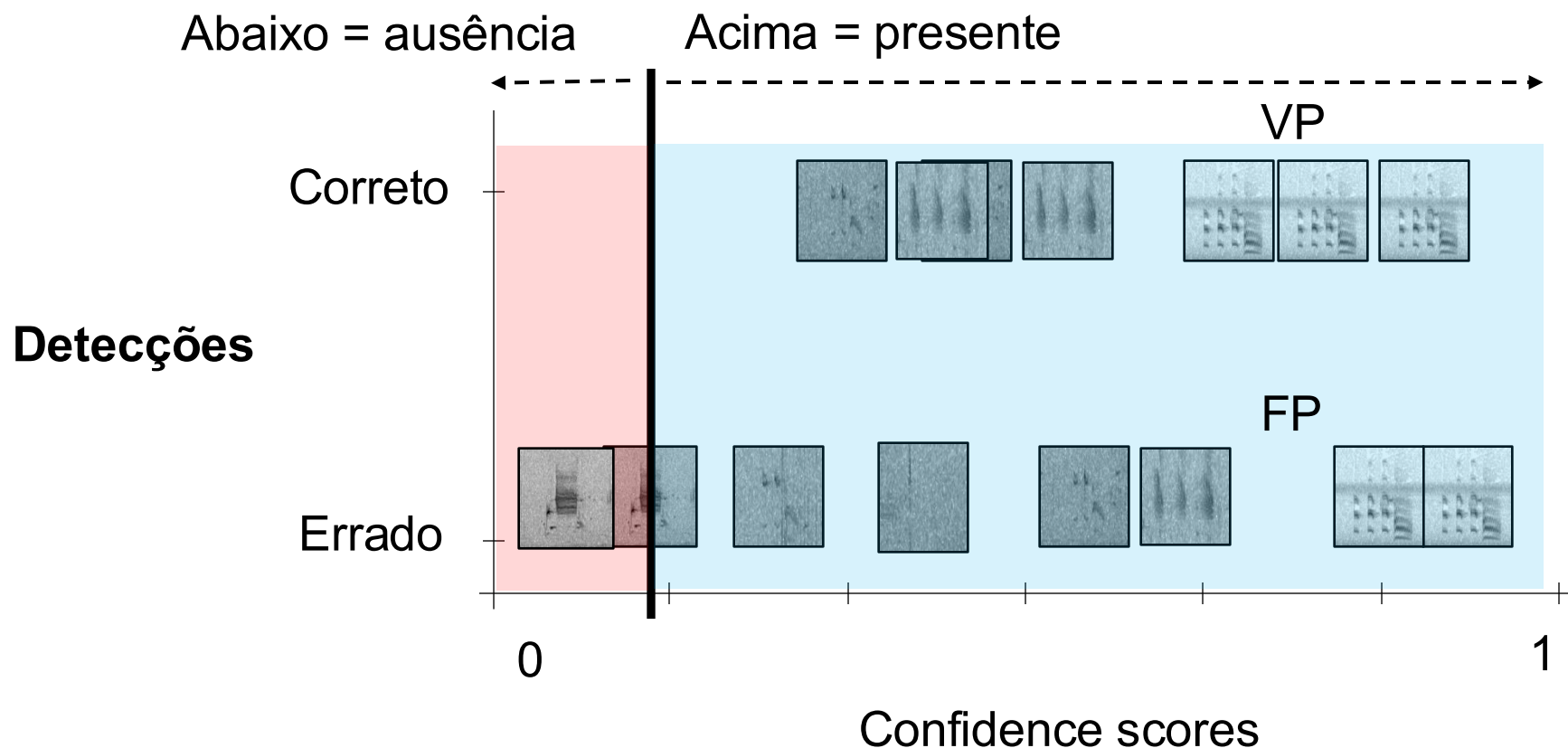
- Critério de corte para classificação: Acima do valor = 1; Abaixo do valor = 0



## 2.1

## Limiar (threshold): tornando uma predição em detecção

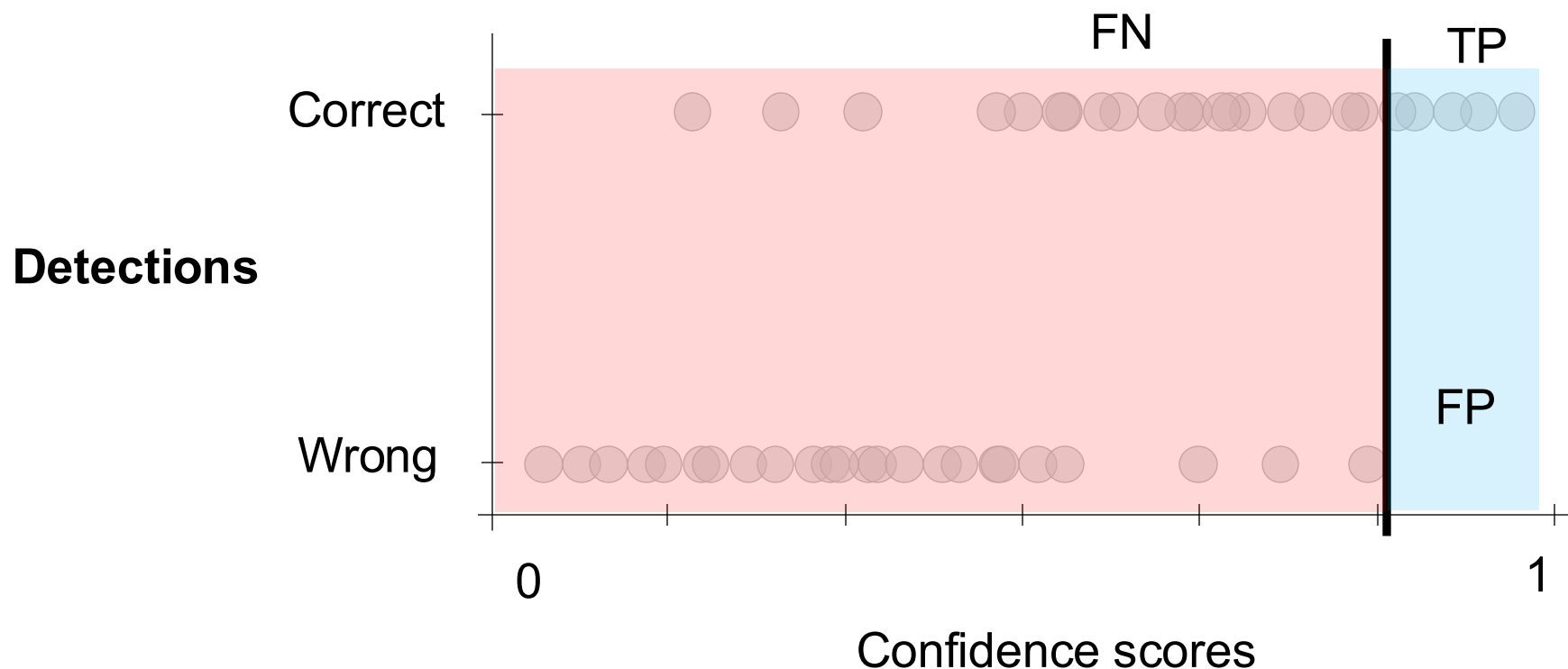
- Critério de corte para classificação: Acima do valor = 1; Abaixo do valor = 0
- O que acontece com a quantidade de VP e FP com limiar baixo?



## 2.1

## Limiar (threshold): tornando uma predição em detecção

- Critério de corte para classificação: Acima do valor = 1; Abaixo do valor = 0
- O que acontece com a quantidade de VP e FP com limiar baixo? E alto?

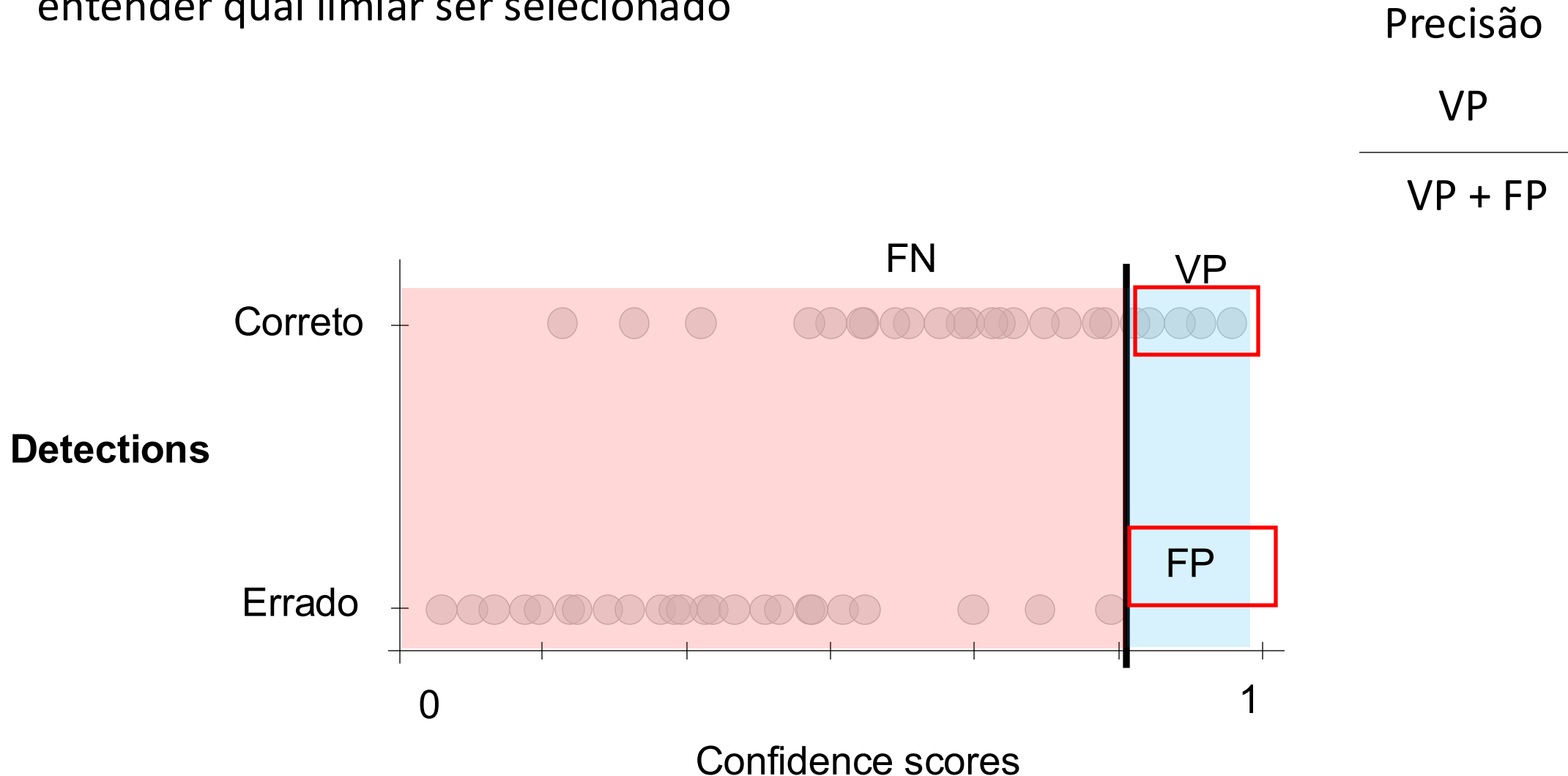




## 2.1

## Limiar (threshold): tornando uma predição em detecção

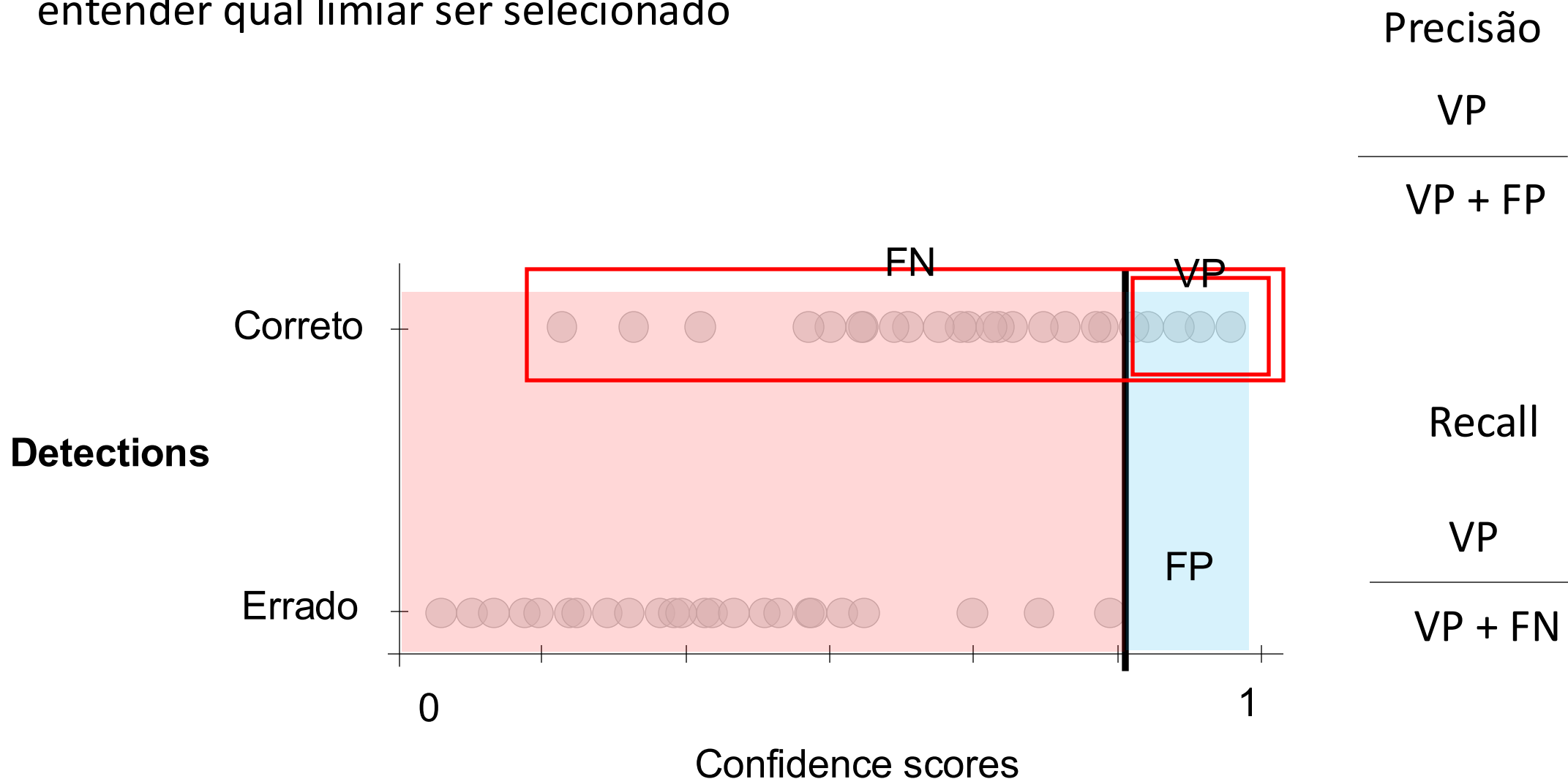
- Precisão e revocação (recall): devem ser analisados no conjunto teste para entender qual limiar ser selecionado



2.1

## Limiar (threshold): tornando uma predição em detecção

- Precisão e revocação (recall): devem ser analisados no conjunto teste para entender qual limiar ser selecionado



**Tutorial!**