

Redução de Custo Experimental na Geração e Avaliação de Dados Sintéticos para Detecção de Malware Android

Lucas Ferreira

Mestrando em Engenharia de Software — PPGES/UNIPAMPA

[tune3-hpo.github.io](https://github.com/tune3-hpo)

Orientador: Prof. Dr. Diego Kreutz

Coorientador: Prof. Dr. Dionathan Schmidt



Dataset	Year	Samples	Features	Benign	Malware
MH100k	2023	101.934	24.833	92.134	9.800
KronoDroid Real	2021	78.137	286	36.755	41.382
KronoDroid Emulator	2021	63.991	276	35.246	28.745
AndroCrawl	2013	96.744	141	86.574	10.170
Android Permissions	2018	26.864	151	9.077	17.787
DREBIN-215	2018	15.031	215	9.476	5.555
DefenseDroid PRS	2021	11.975	2.877	5.975	6.000
DD API Katz	2021	10.476	6.002	5.222	5.254
DD API Degree	2021	10.476	6.002	5.222	5.254
DD API Closeness	2021	10.476	4.274	5.222	5.254
Adroit	2017	11.476	166	8.058	3.418

Dataset	Year	Samples	Features	Benign	Malware
MH100k	2023	101.934	24.833	92.134	9.800
KronoDroid Real	2021	78.137	286	36.755	41.382
KronoDroid Emulator	2021	63.991	276	35.246	28.745
AndroCrawl	2013	96.744	141	86.574	10.170
Android Permissions	2018	26.864	151	9.077	17.787
DREBIN-215	2018	15.031	215	9.476	5.555
DefenseDroid PRS	2021	11.975	2.877	5.975	6.000
DD API Katz	2021	10.476	6.002	5.222	5.254
DD API Degree	2021	10.476	6.002	5.222	5.254
DD API Closeness	2021	10.476	4.274	5.222	5.254
Adroit	2017	11.476	166	8.058	3.418

Dataset	Year	Samples	Features	Benign	Malware
MH100k	2023	101.934	24.833	92.134	9.800
KronoDroid Real	2021	78.137	286	36.755	41.382
KronoDroid Emulator	2021	63.991	276	35.246	28.745
AndroCrawl	2013	96.744	141	86.574	10.170
Android Permissions	2018	26.864	151	9.077	17.787
DREBIN-215	2018	15.031	215	9.476	5.555
DefenseDroid PRS	2021	11.975	2.877	5.975	6.000
DD API Katz	2021	10.476	6.002	5.222	5.254
DD API Degree	2021	10.476	6.002	5.222	5.254
DD API Closeness	2021	10.476	4.274	5.222	5.254
Adroit	2017	11.476	166	8.058	3.418

Dataset	Year	Samples	Features	Benign	Malware
MH100k	2023	101.934	24.833	92.134	9.800
KronoDroid Real	2021	78.137	286	36.755	41.382
KronoDroid Emulator	2021	63.991	276	35.246	28.745
AndroCrawl	2013	96.744	141	86.574	10.170
Android Permissions	2018	26.864	151	9.077	17.787
DREBIN-215	2018	15.031	215	9.476	5.555
DefenseDroid PRS	2021	11.975	2.877	5.975	6.000
DD API Katz	2021	10.476	6.002	5.222	5.254
DD API Degree	2021	10.476	6.002	5.222	5.254
DD API Closeness	2021	10.476	4.274	5.222	5.254
Adroit	2017	11.476	166	8.058	3.418

Dataset**Dimensão (Samples × Features)****Exec (horas)**

MH100k

~2,53 bilhões

28.583,6

KronoDroid Real

~22,3 milhões

280,6

KronoDroid Emulator

~18,3 milhões

234,9

AndroCrawl

~13,6 milhões

182,7

Android Permissions

~4,0 milhões

73,8

DREBIN-215

~3,2 milhões

64,4

DefenseDroid PRS

~34,4 milhões

416,4

DD API Katz

~62,9 milhões

736,9

DD API Degree

~62,9 milhões

736,9

DD API Closeness

~44,8 milhões

532,8

Adroit

~1,9 milhões

49,4

**Alta
dimensionalidade
aumenta o custo
computacional**

Dataset

Dimensão (Samples × Features)

Exec (horas)

MH100k

~2,53 bilhões

28.583,6

KronoDroid Real

~22,3 mil

280,6

KronoDroid Emulator

~18,3 m

234,9

AndroCrawl

~13,6 m

182,7

Android Permissions

~4,0 milhões

73,8

DREBIN-215

~3,2 milhões

64,4

DefenseDroid PRS

~34,4 milhões

416,4

DD API Katz

~62,9 milhões

736,9

DD API Degree

~62,9 milhões

736,9

DD API Closeness

~44,8 milhões

532,8

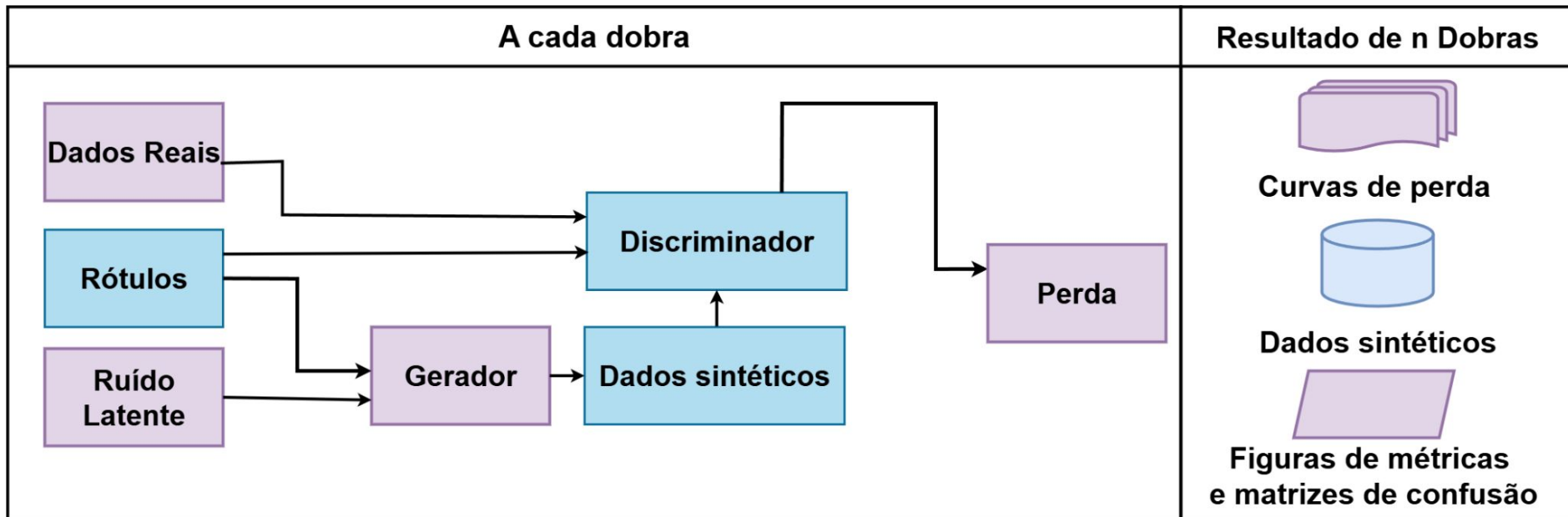
Adroit

~1,9 milhões

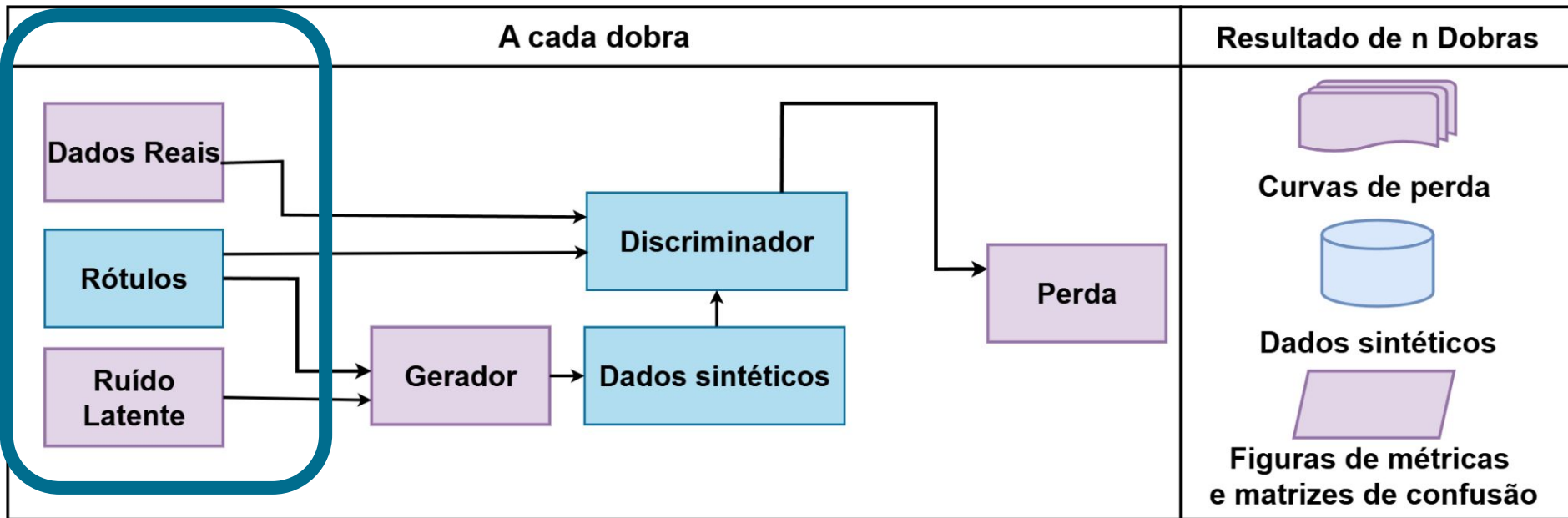
49,4

**~ 3 anos de execução
na MalSynGen**

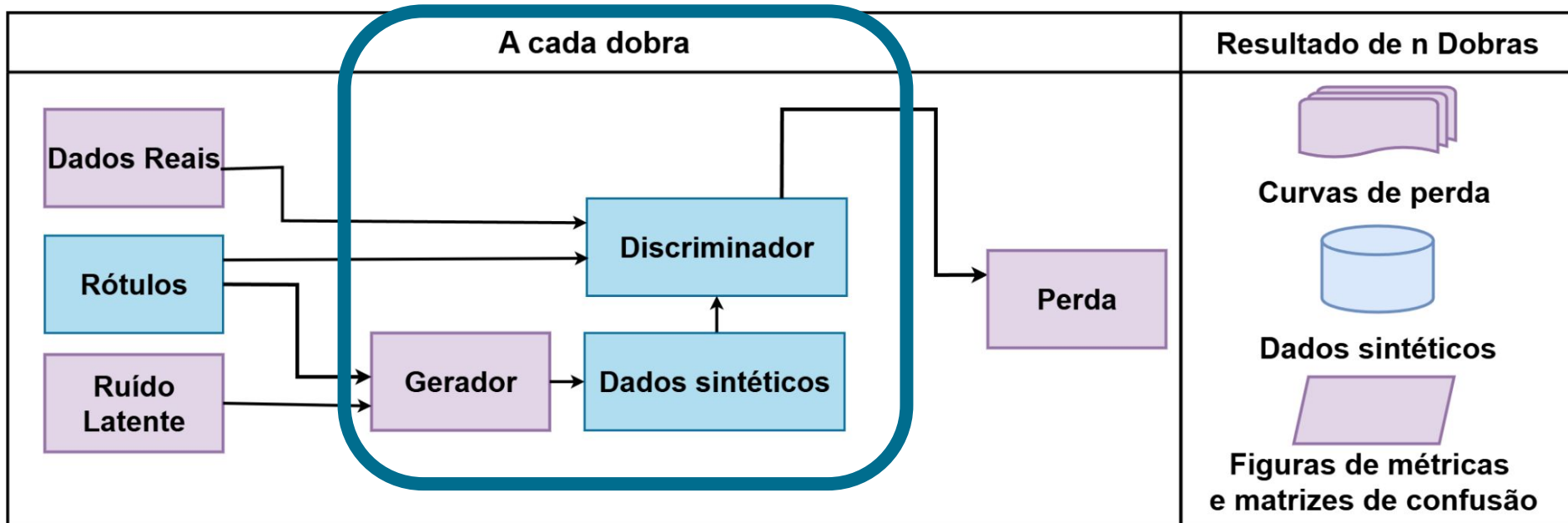
MalSynGen



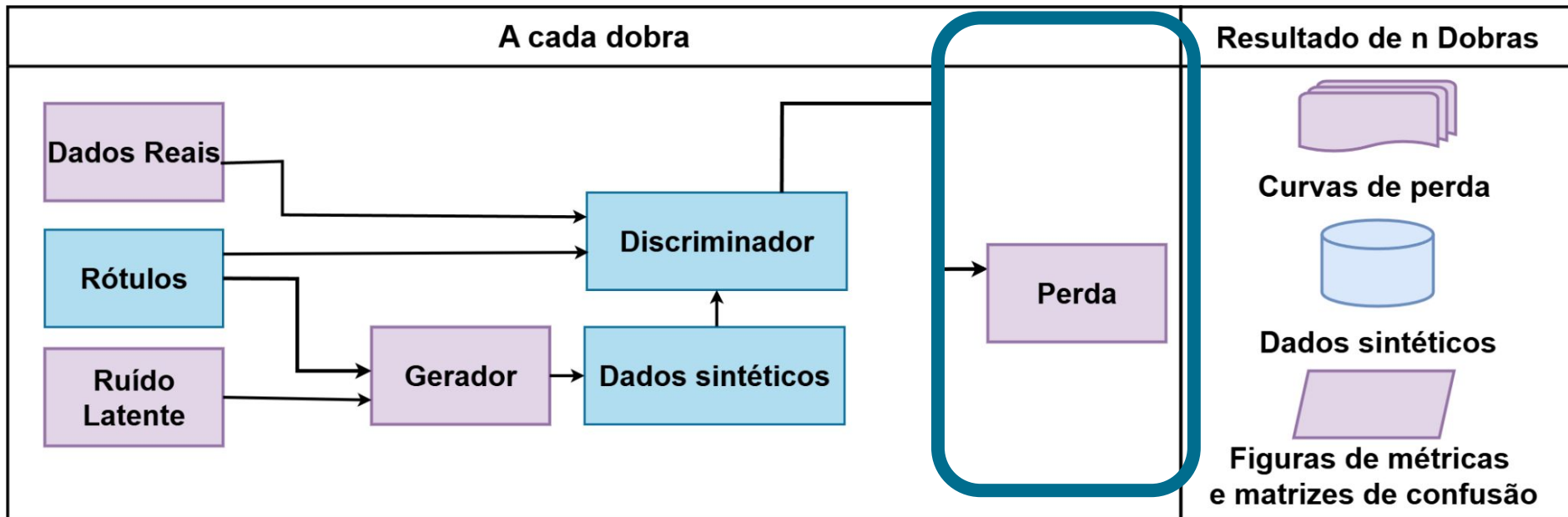
MalSynGen



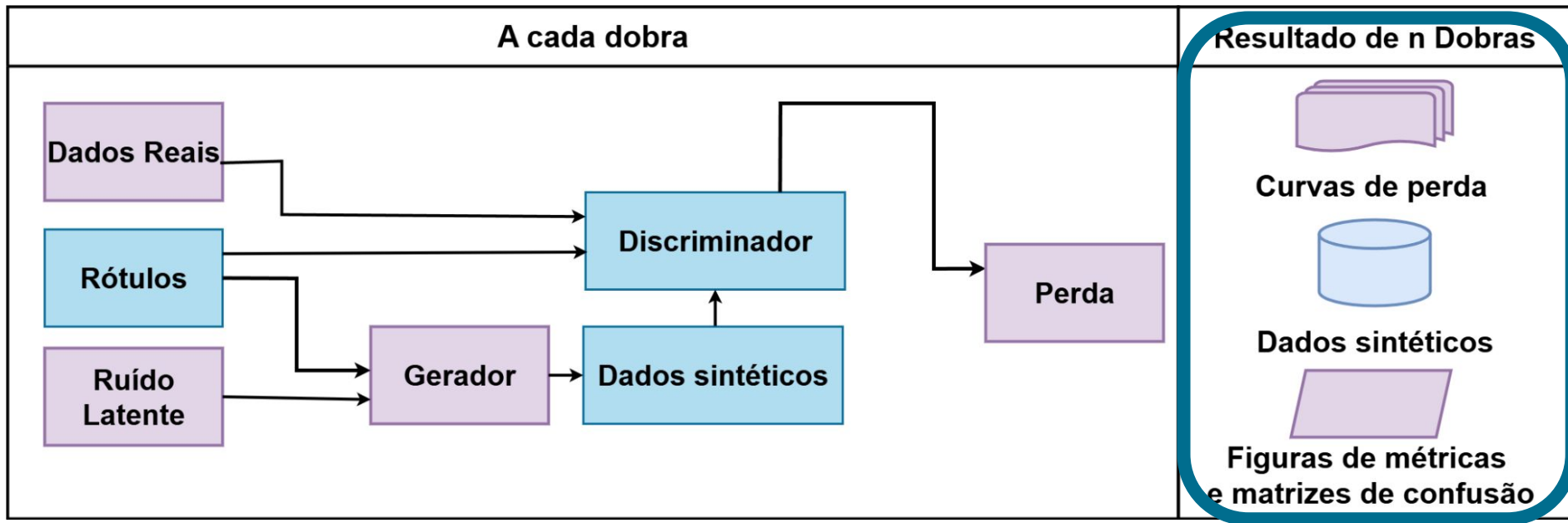
MalSynGen



MalSynGen



MalSynGen



Hiperparâmetros

Hiperparâmetros	Valores
K_Fold	3 a 10
Densidade de Neurônios por camada	8 - 4096
Número de épocas	1-5000
Dropout D/G	0.00 - 1.00
Algoritmo de treinamento	Adam, Nadam, Adamax, ...
Função de ativação	ReLU, LeakyReLU, PReLU,...

Hiperparâmetros

**Hiperparâmetros
identificados na
literatura e
suportados pelo
MalSynGen**

Valores

3 a 10

8 - 4096

1-5000

0.00 - 1.00

Adam, Nadam, Adamax, ...

ReLU, LeakyReLU, PReLU,...

Hiperparâmetros

**+ - 2 bilhões de
anos para
executar um
único dataset**

Valores

3 a 10

8 - 4096

1-5000

0.00 - 1.00

Adam, Nadam, Adamax, ...

ReLU, LeakyReLU, PReLU,...

Questão central: Como reduzir o custo experimental da otimização de hiperparâmetros, preservando o desempenho preditivo?

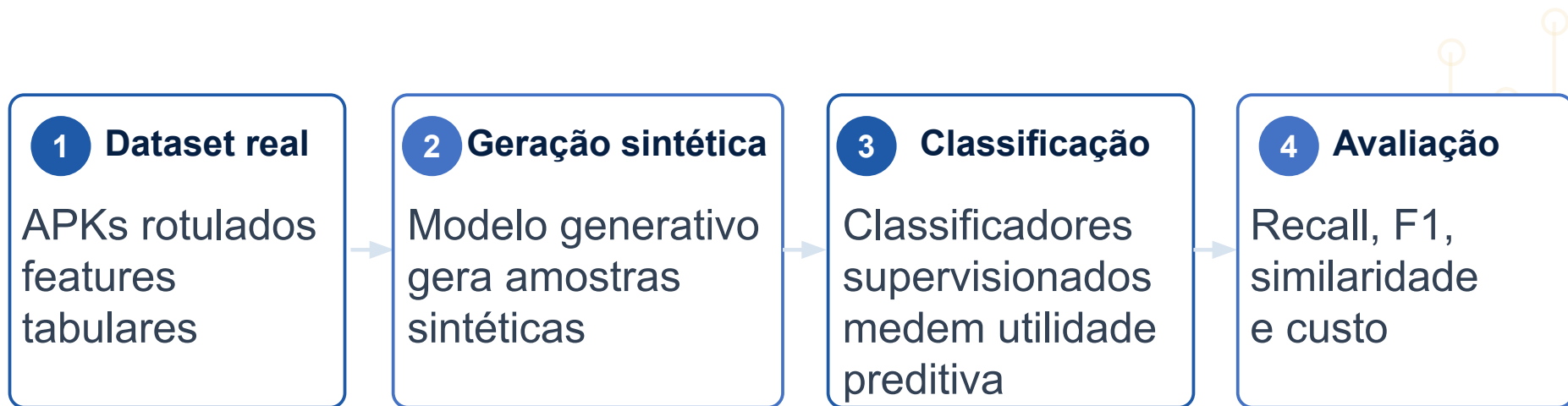


ARTIGO 1 / 4

Otimização de hiperparâmetros da DroidAugmentor para geração de dados sintéticos de malware Android

Artigo completo · ERRC/WRSeg 2024 · com A. Nogueira, A. L. Gomes, D. Kreutz, R. Mansilha

Dados sintéticos para detecção de malware Android



Ponto de partida

Dados sintéticos podem apoiar a detecção, mas sua utilidade depende das decisões tomadas antes da avaliação.

Hiperparâmetros

Como os hiperparâmetros em modelos generativos afetam a qualidade dos dados sintéticos e o custo computacional?

Entrada

10 datasets
tabulares de
malware Android

Processo

cGAN com
variação de
camadas, épocas
e dropout

Saída

métricas de
utilidade,
similaridade e
custo

Dataset	Features	Samples
defensedroid_apicalls_katz	200	10444
defensedroid_apicalls_degree	200	10444
defensedroid_apicalls_closeness	200	10444
kronodroid_emulator	200	20000
kronodroid_real_device	200	20000
adroit	118	6836
defensedroid_prs	200	11950
drebin215	200	11110
android_permissions	148	18154
PPGES · UNIPAMPA androcrawl	136	20340

Dataset	Features	Samples
defensedroid_apicalls_katz	200	10444
defensedroid_apicalls_degree	200	10444
defensedroid_apicalls_closeness	200	10444
kronodroid_emulator	200	20000
kronodroid_real_device	200	20000
adroit	118	6836
defensedroid_prs	200	11950
drebin215	200	11110
android_permissions	148	18154
PPGES · UNIPAMPA androcrawl	136	20340

Dataset	Features	Samples
defensedroid_apicalls_katz	200	10444
defensedroid_apicalls_degree	200	10444
defensedroid_apicalls_closeness	200	10444
kronodroid_emulator	200	20000
kronodroid_real_device	200	20000
adroit	118	6836
defensedroid_prs	200	11950
drebin215	200	11110
android_permissions	148	18154
PPGES · UNIPAMPA androcrawl	136	20340

Análise de hiperparâmetros

Impacto dos hiperparâmetros de uma cGAN na geração (10 datasets)

10

datasets
avaliados

3

classificadores
de utilidade
RF, KNN e DT

3 blocos

utilidade,
similaridade e
custo

Hiperparâmetros avaliados

Parâmetro	Valores
Épocas	1000, 2000, 5000
Neurônios por camada	64, 256, 512, 1024, 2048, 4096
Dropout D	0.05, 0.1, 0.2, 0.3, 0.4
Dropout G	0.0025, 0.05, 0.1, 0.15, 0.2
Otimização / Ativação	Adam / LeakyReLU

Resultados

Hiperparâmetros mudam, ao mesmo tempo, qualidade e custo

Neurônios por camada

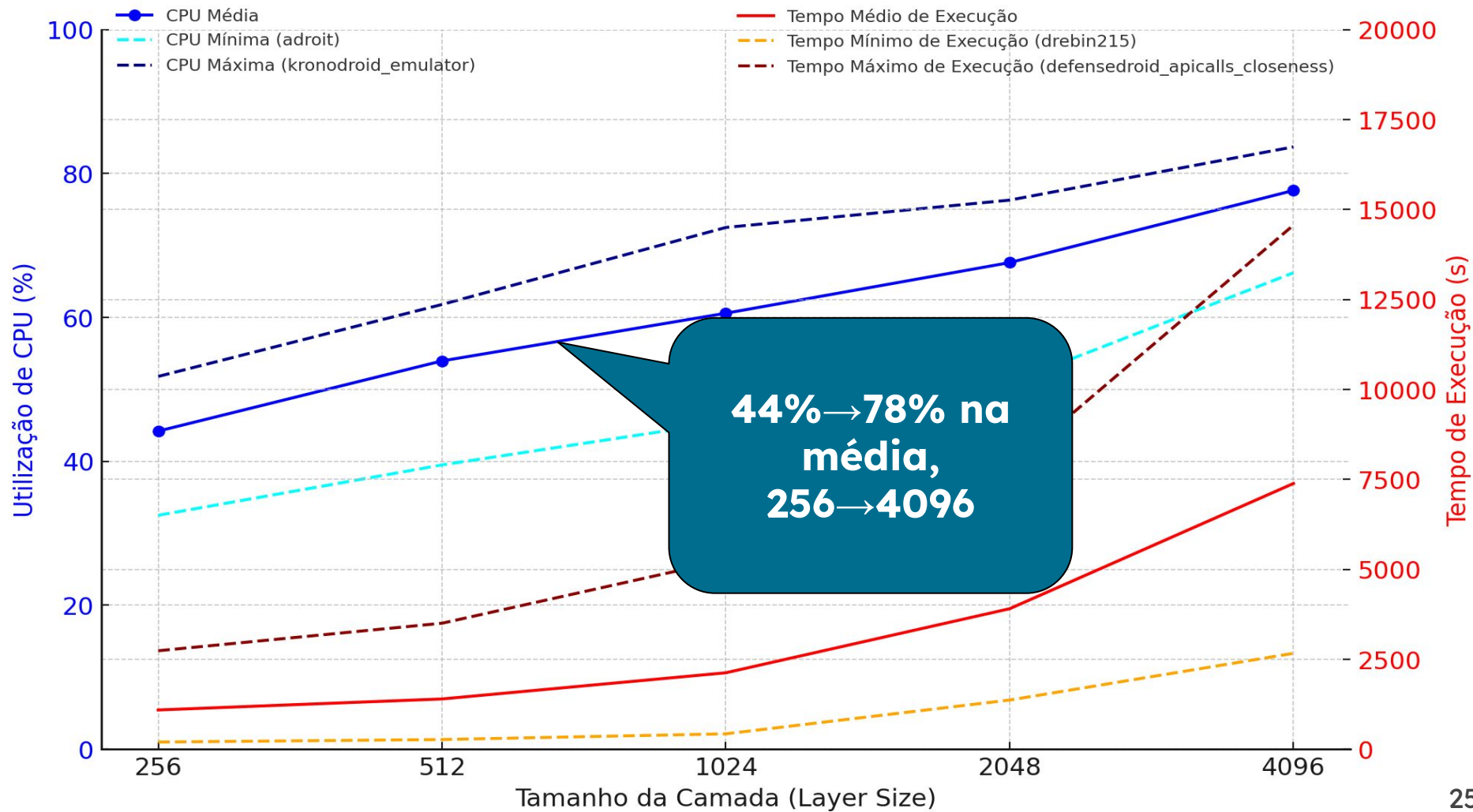
→ aumenta o uso de CPU e o tempo de execução

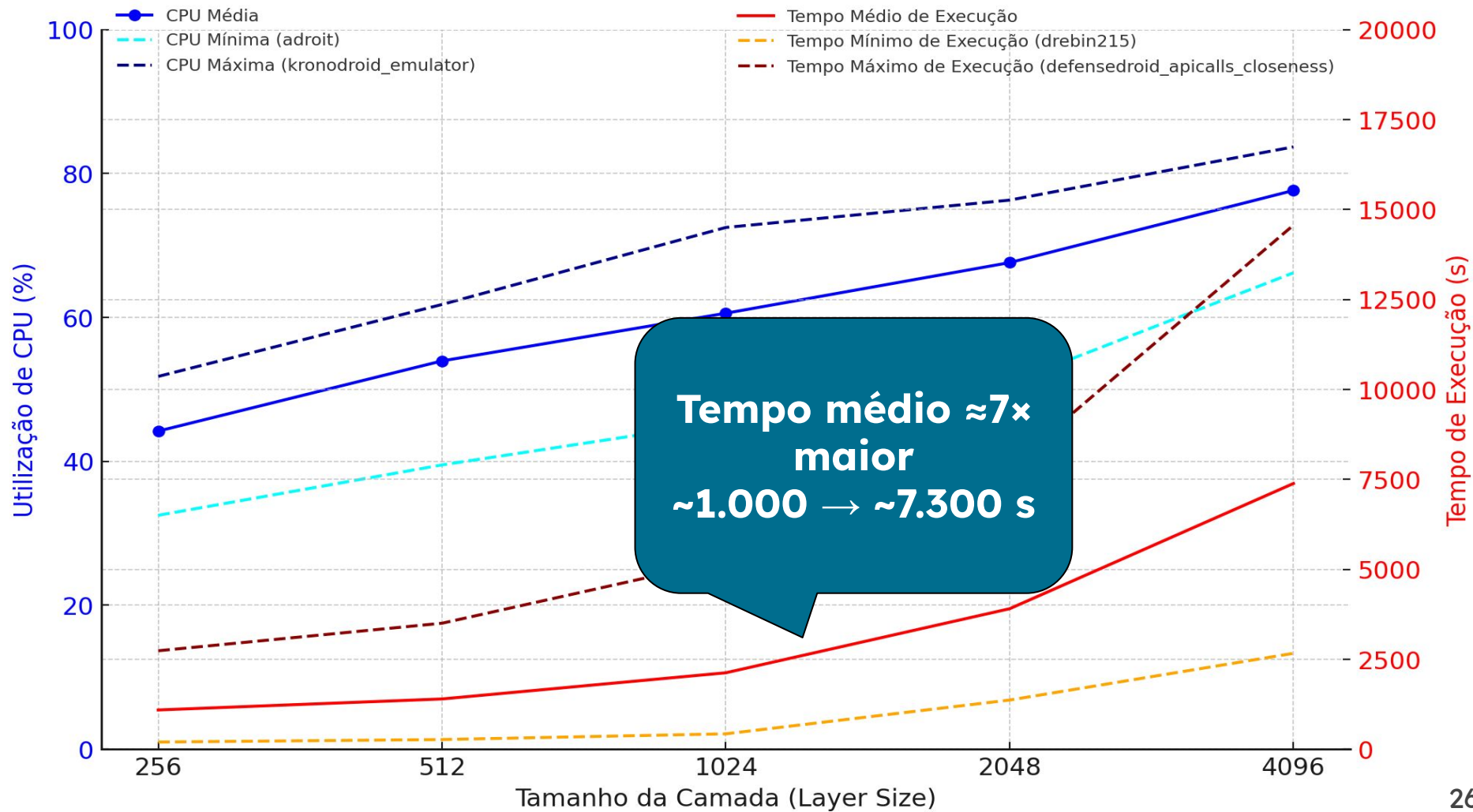
Número de épocas

→ aumenta o consumo de memória; CPU permanece estável

Configuração ótima

→ muda a cada dataset — não existe valor único





Resultados

Hiperparâmetros mudam, ao mesmo tempo

**Aumentar as épocas
não garante melhoria
consistente e
aumenta o custo
experimental**

Densidade de camadas

→ aumenta o tempo de execução

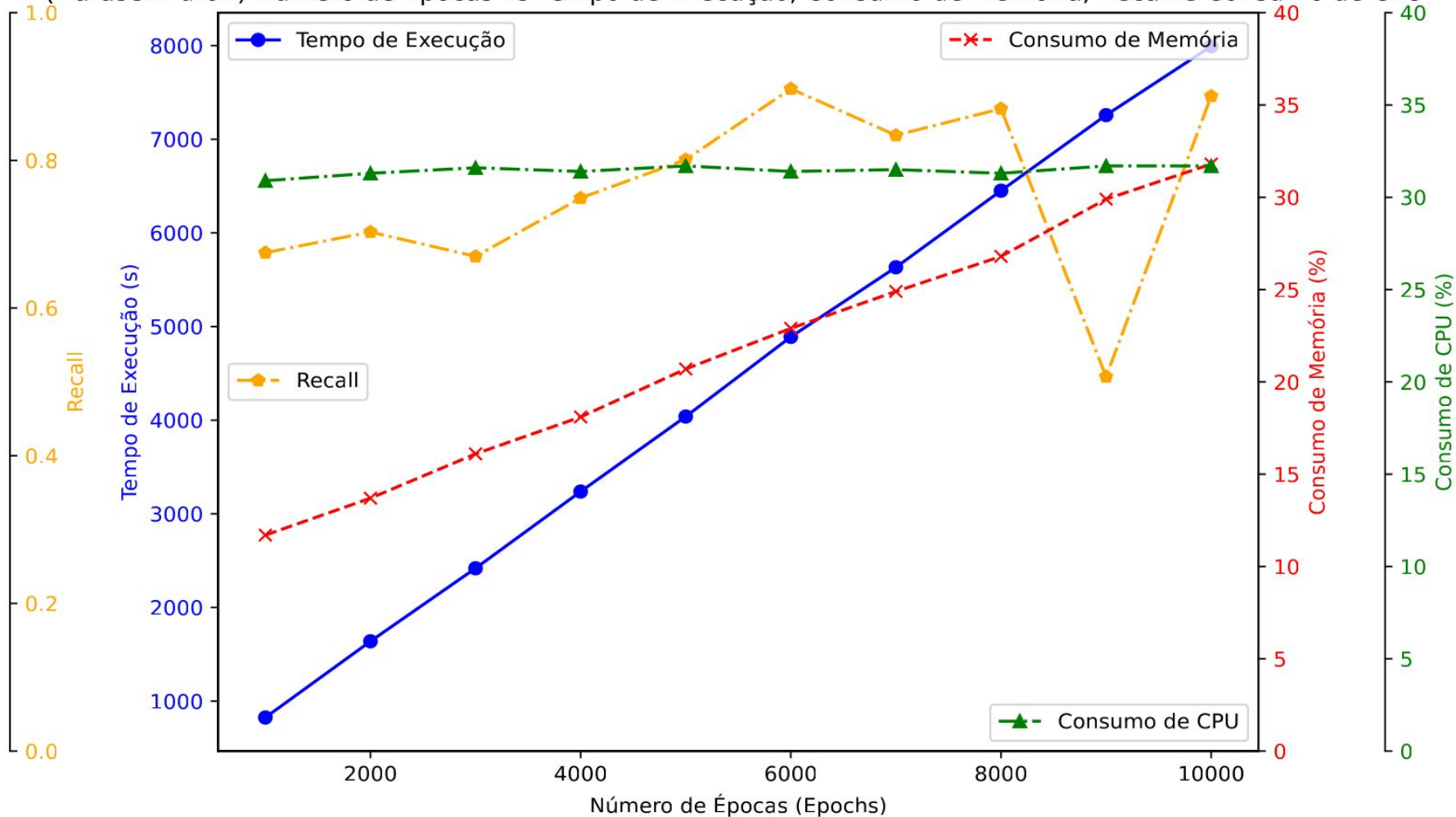
Número de épocas

→ aumenta o consumo de memória; CPU permanece estável

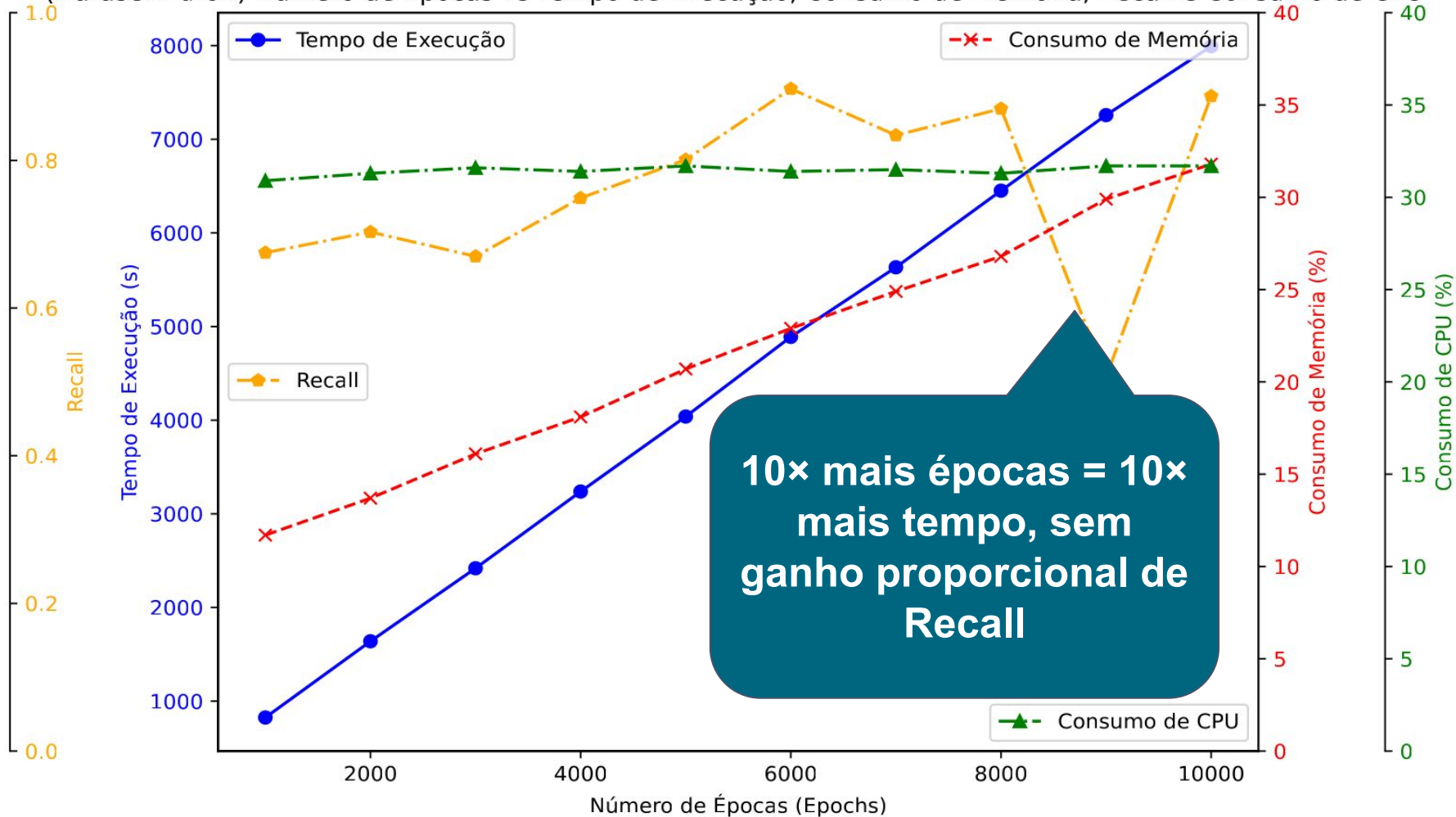
Configuração ótima

→ muda a cada dataset — não existe valor único

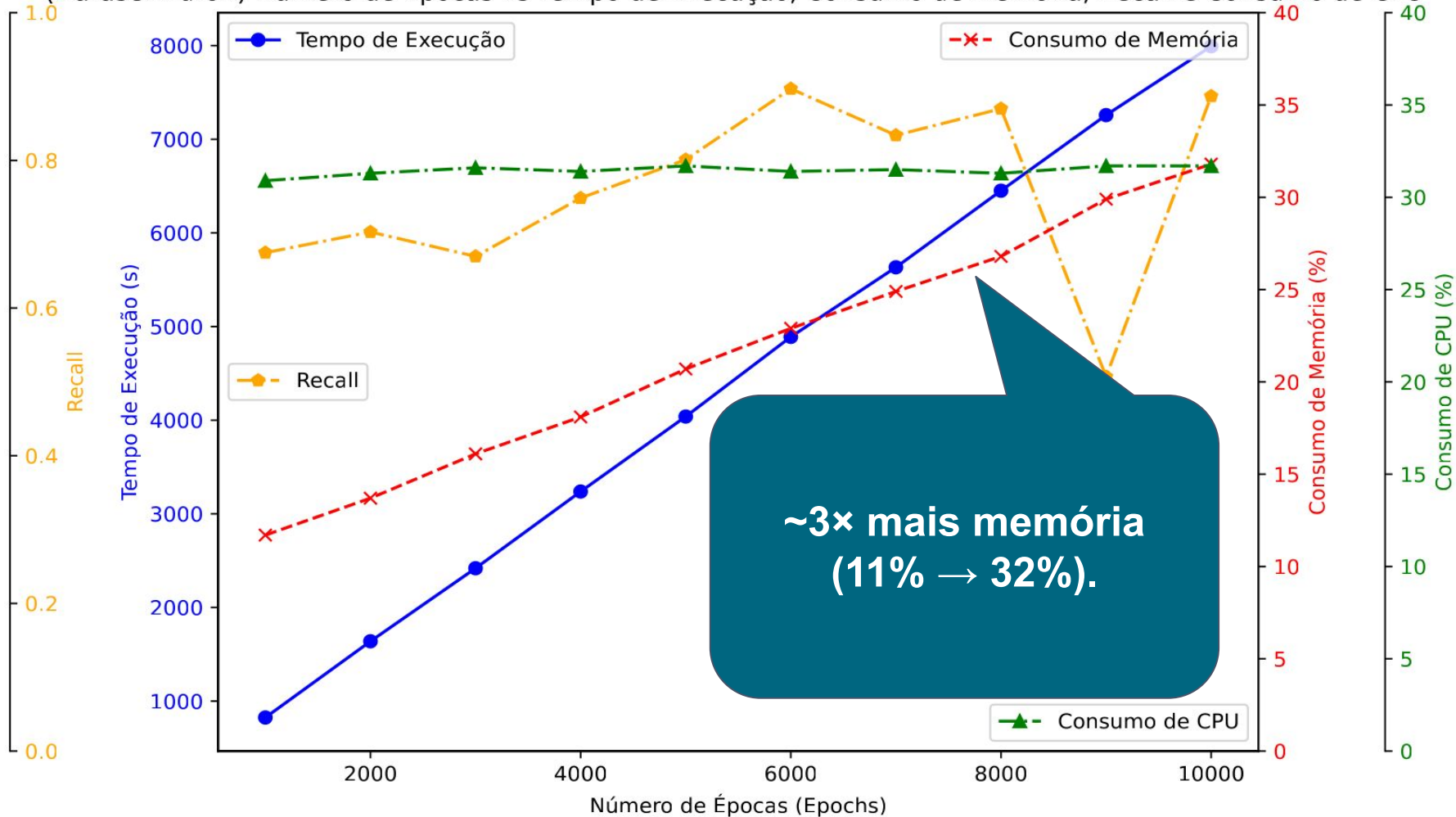
(Dataset Adroit) Número de Épocas vs Tempo de Execução, Consumo de Memória, Recall e Consumo de CPU



(Dataset Adroit) Número de Épocas vs Tempo de Execução, Consumo de Memória, Recall e Consumo de CPU



(Dataset Adroit) Número de Épocas vs Tempo de Execução, Consumo de Memória, Recall e Consumo de CPU



Resultados

Hiperparâmetros mudam, ao mesmo tempo

Densidade de camadas

→ aumenta o tempo de execução

Número de épocas

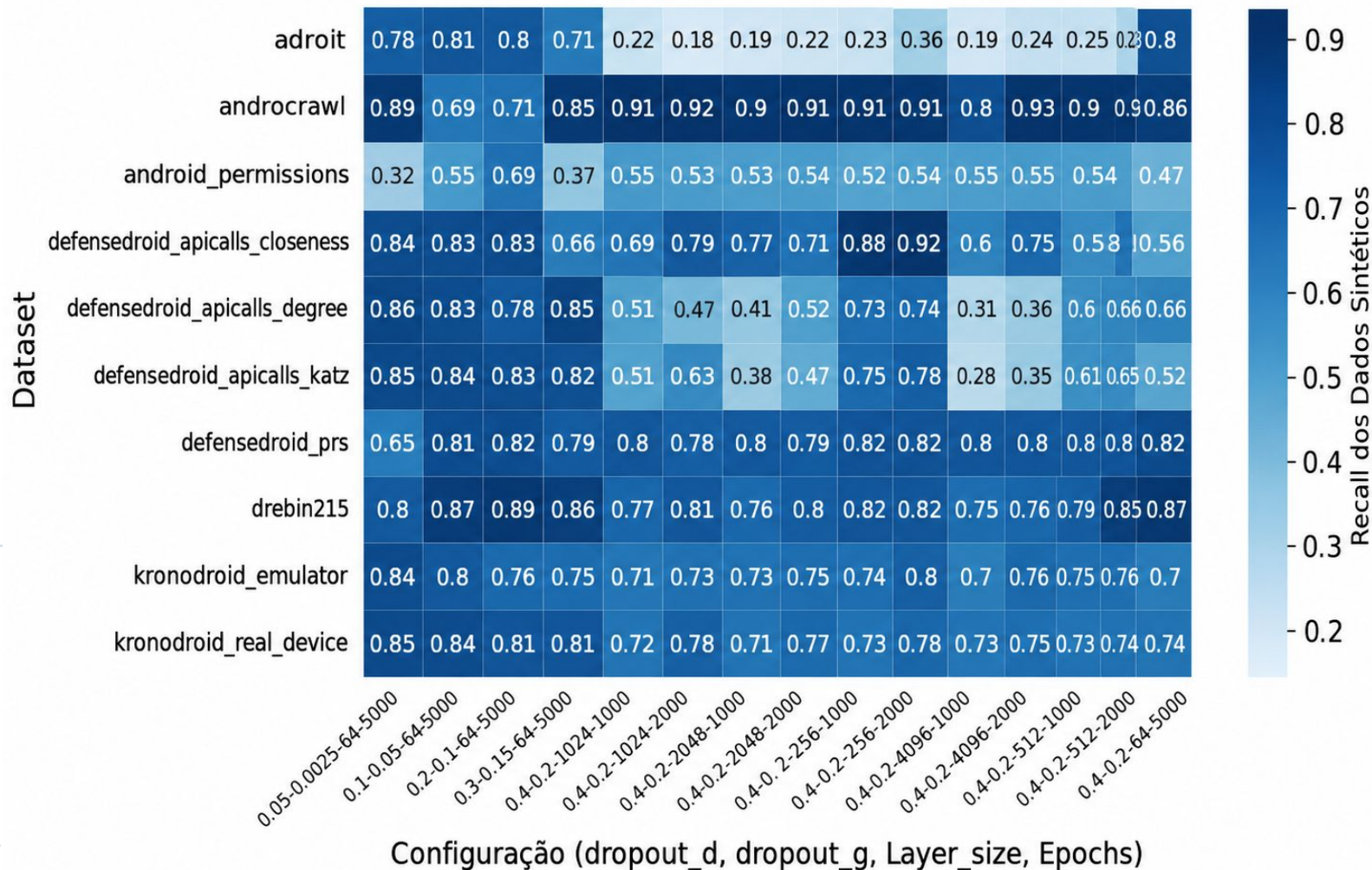
→ aumenta o tempo de execução e permanece estável

Configuração ótima

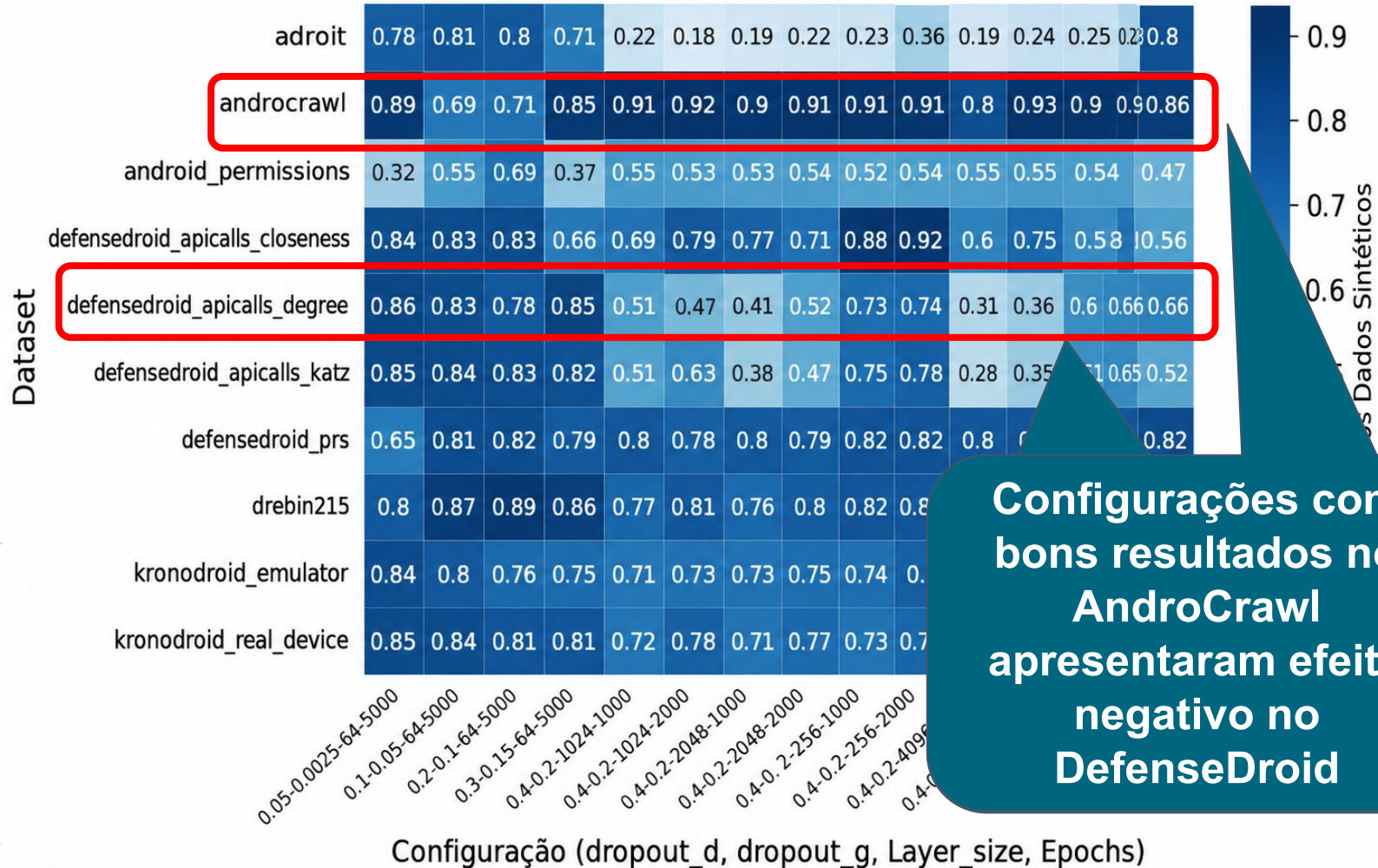
→ muda a cada dataset — não existe valor único

**Não existe uma
configuração ideal
que atenda todos os
datasets**

Heatmap de Recall dos Dados Sintéticos por Configuração e Dataset



Heatmap de Recall dos Dados Sintéticos por Configuração e Dataset



Achado anterior

Hiperparâmetros alteram qualidade, estabilidade e custo.



Pergunta

Como priorizar configurações promissoras e evitar as inviáveis



Resposta

Buscar uma técnica eficiente de HPO

O próximo passo é transformar esses achados em uma estratégia de busca, voltada a reduzir combinações pouco promissoras e orientar a exploração experimental.

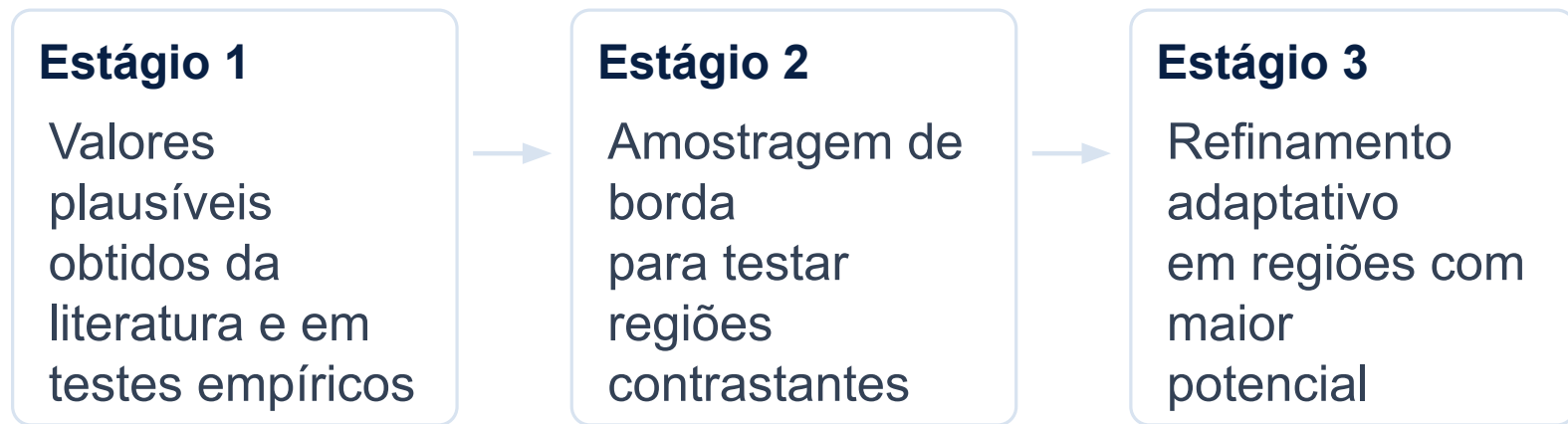
ARTIGO 2 / 4

Tune3: busca multiestágio de hiperparâmetros com redução de custo e maior estabilidade

Artigo completo · ERRC/WRSeg 2025 · com A. Diniz, D. Kreutz, D. Schmidt, R. Mansilha

Tune3: Busca multiestágio para redução do espaço experimental

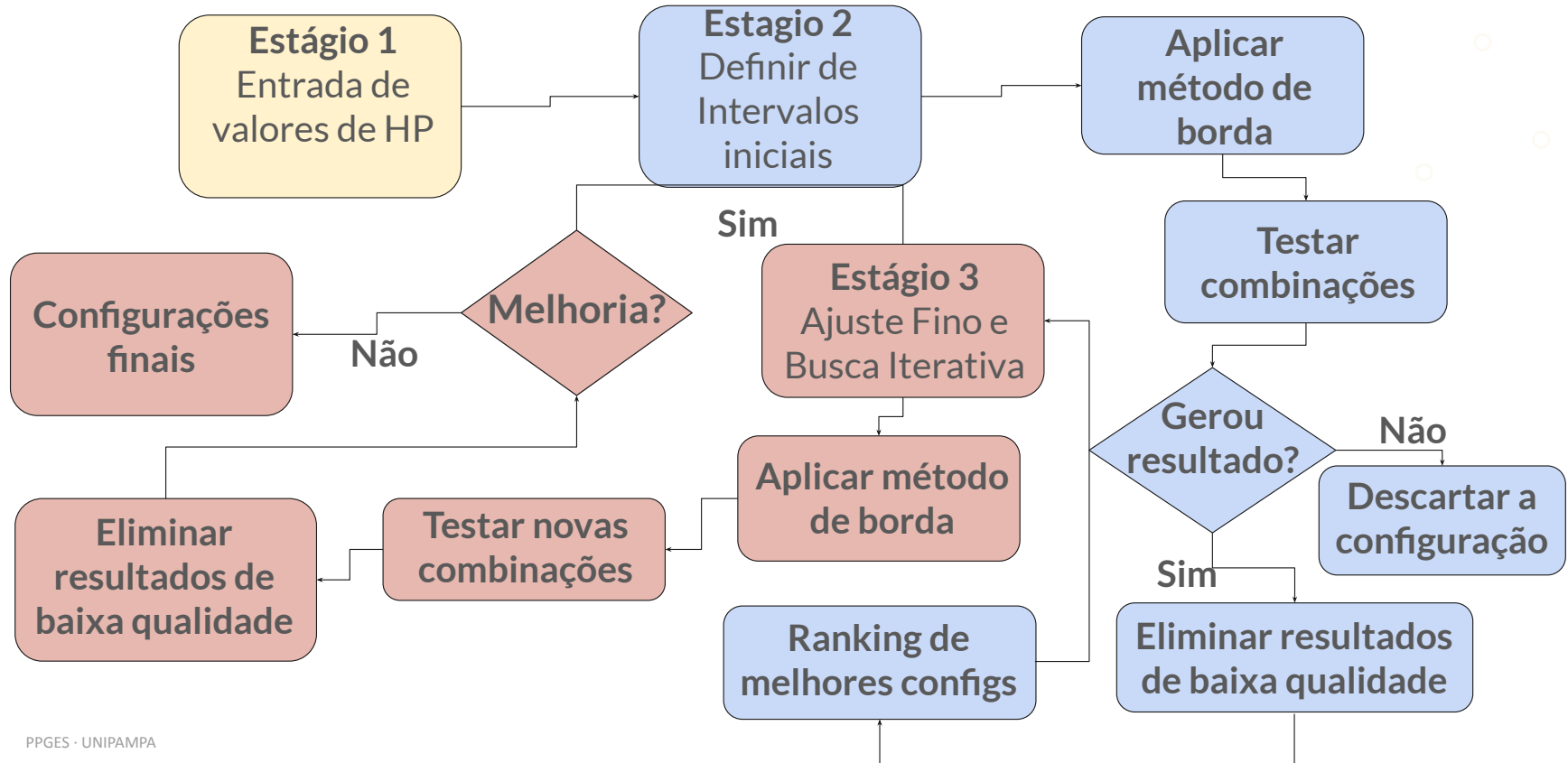
O método reduz o espaço por filtragem progressiva, não por busca exaustiva



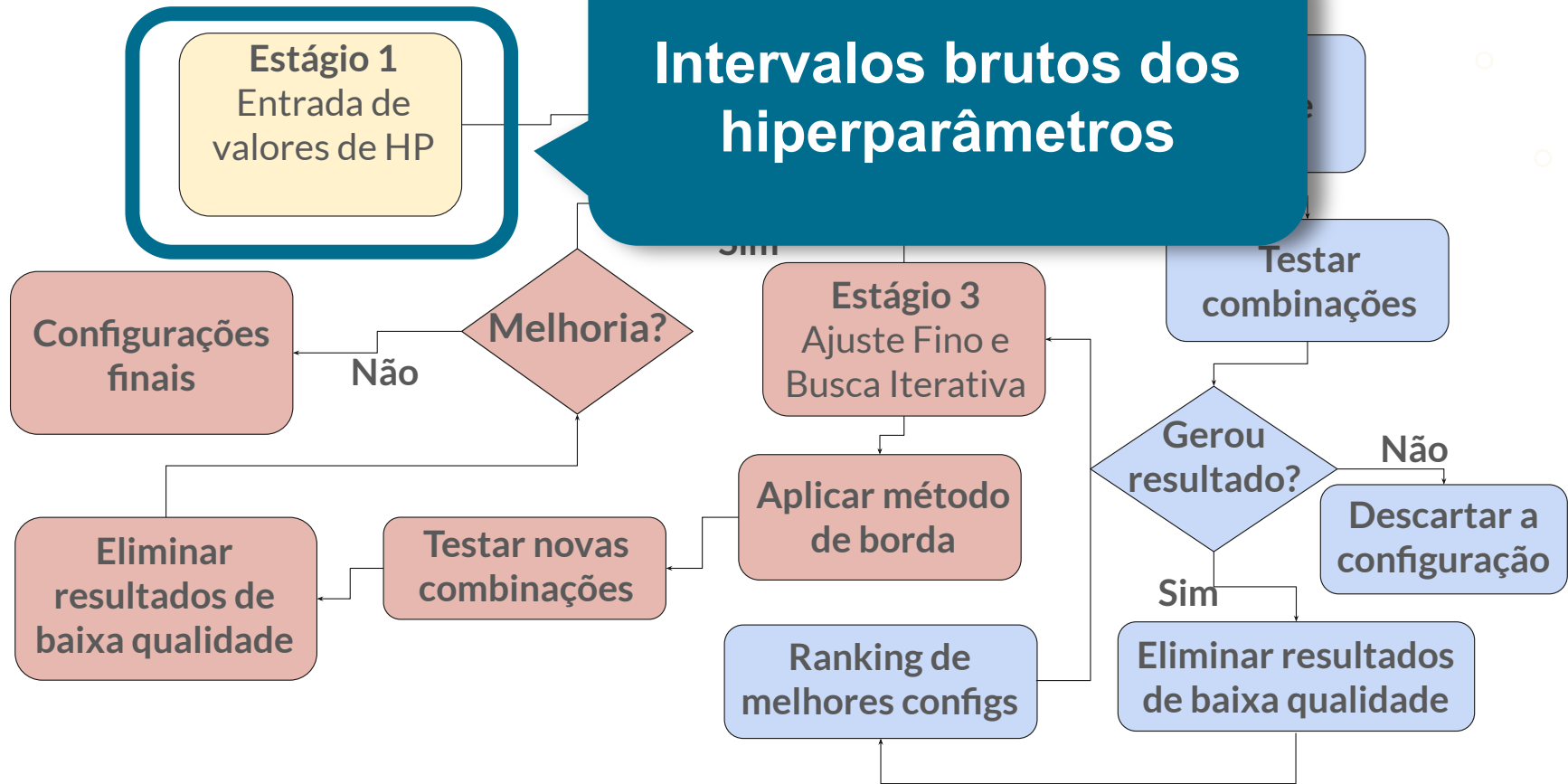
CrITÉrios de decisão usados no processo

- priorização de Recall e F1-score
- remoção de configurações instáveis ou inviáveis

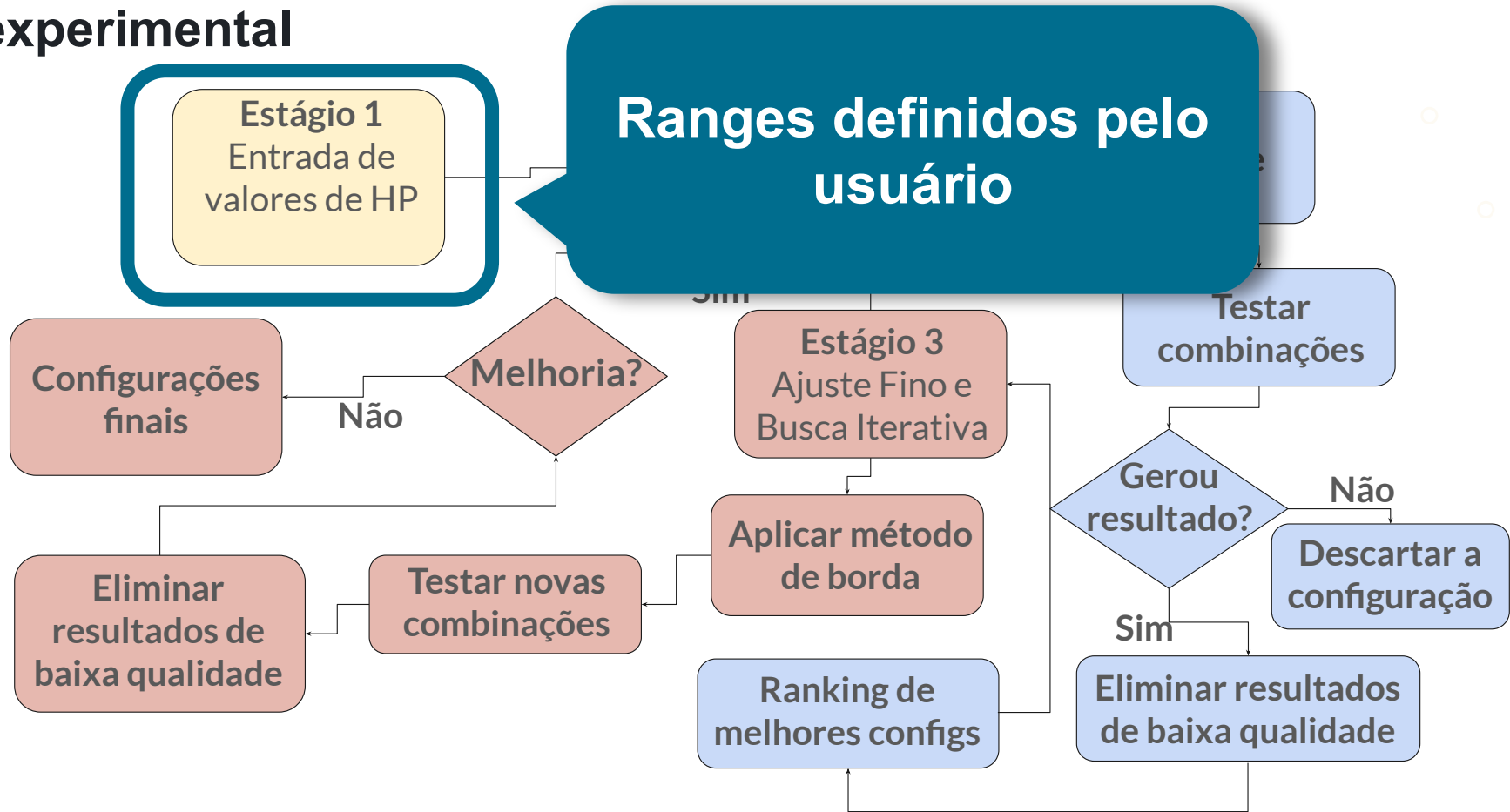
Tune3: Busca multiestágio para redução do espaço experimental



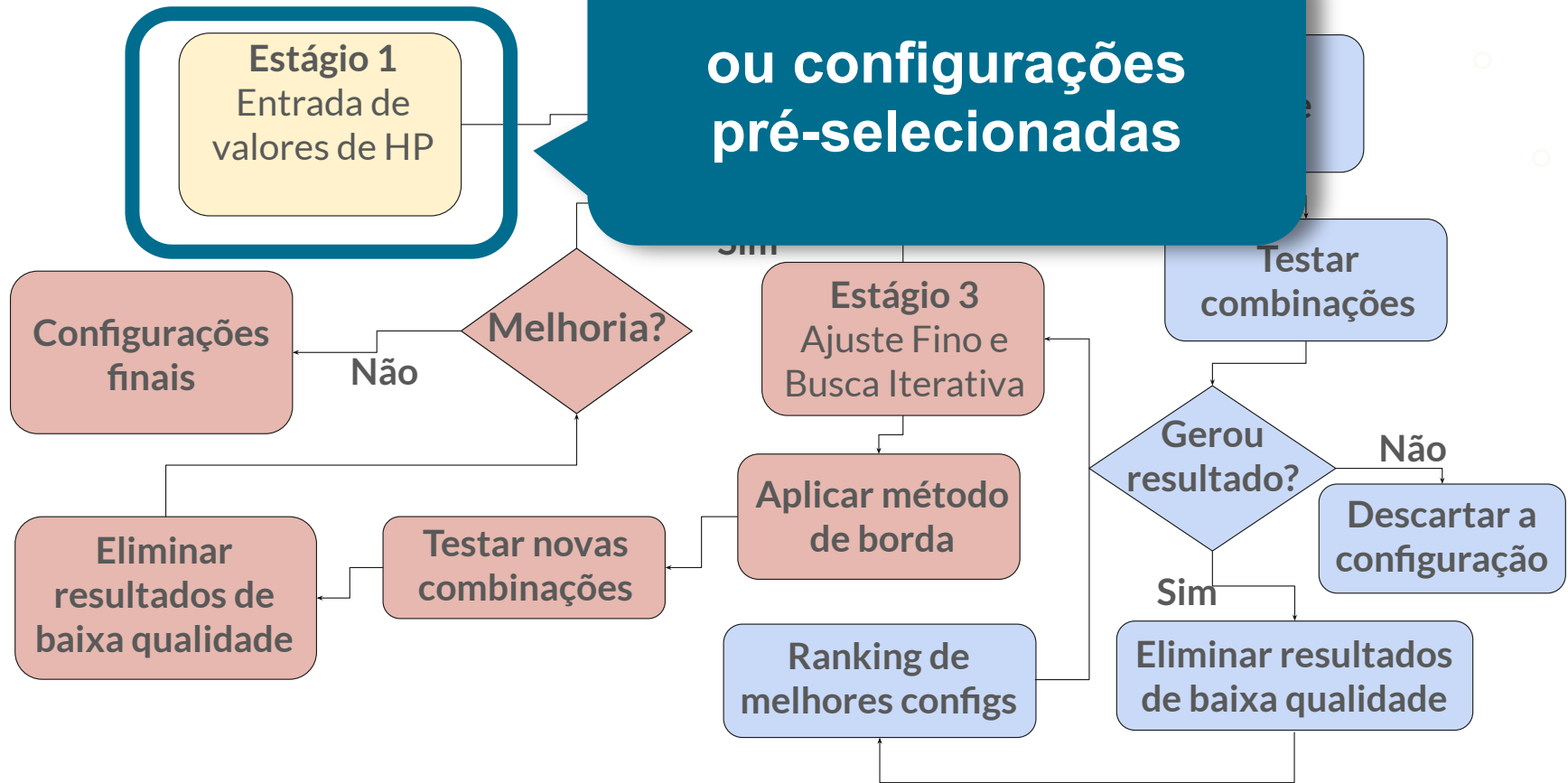
Tune3: Busca multiestágio para redução do espaço experimental



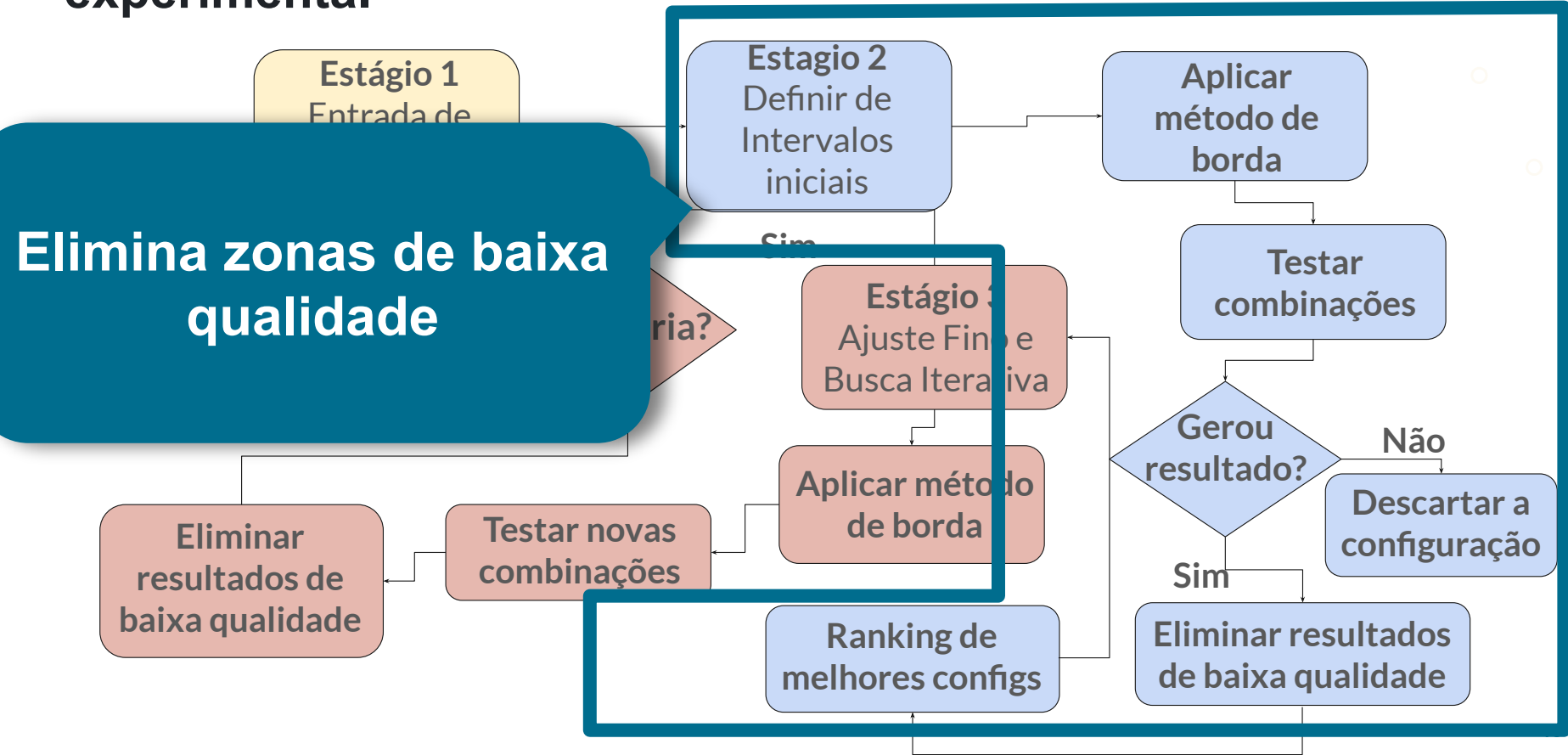
Tune3: Busca multiestágio para redução do espaço experimental



Tune3: Busca multiestágio para redução do espaço experimental

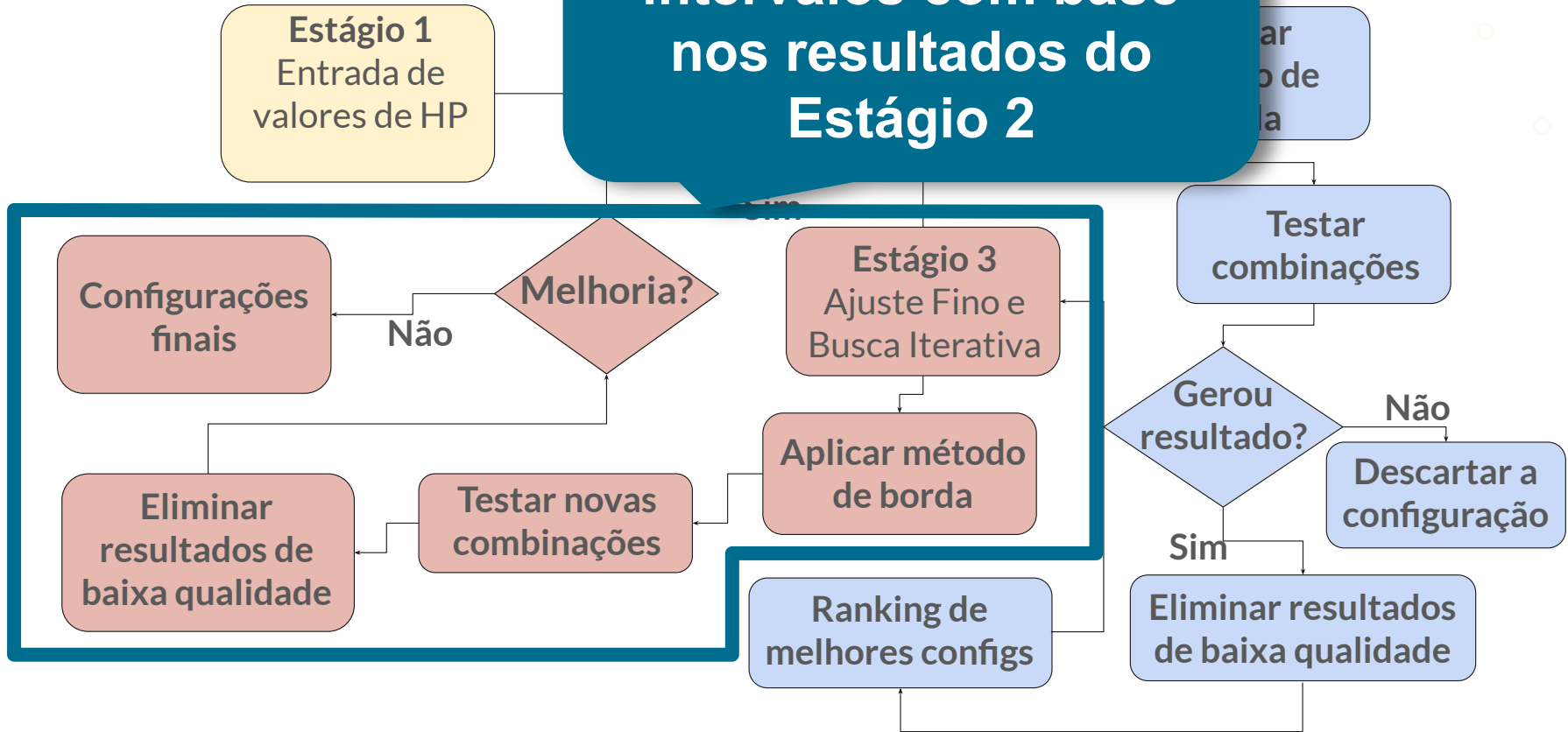


Tune3: Busca multiestágio para redução do espaço experimental



Tune3: Busca multiestágio experimental

Refinamento dos intervalos com base nos resultados do Estágio 2



Datasets utilizados

Dataset	Features	Type Features	Samples
DREBIN-215	215	P, I, A	11110
DefenseDroid APICalls D	200	API Calls	10444

Protocolo experimental

Métodos

Espaço de busca

Execuções

Validação

Métrica principal

Métrica complementar

Tune3 e Random Search

1.125 configurações

até 150 por método e por dataset

5 folds

Recall

F1-score

Hiperparâmetros	Valores
Neurônios por camada G	128 a 1024, com incremento de 64
Neurônios por camada D	128 a 1024, com incremento de 64
Épocas	50, 100, 200, 500 e 1000

Resultados

Dataset	Recall Tune3	Recall Random Search	Dif.	Tempo	Leitura
DREBIN-215	0,96	0,95	+0,01	-34%	Recall equivalente; menor custo
DefenseDroid API Degree	0,92	0,87	+0,05	-44%	Melhor Recall; menor custo

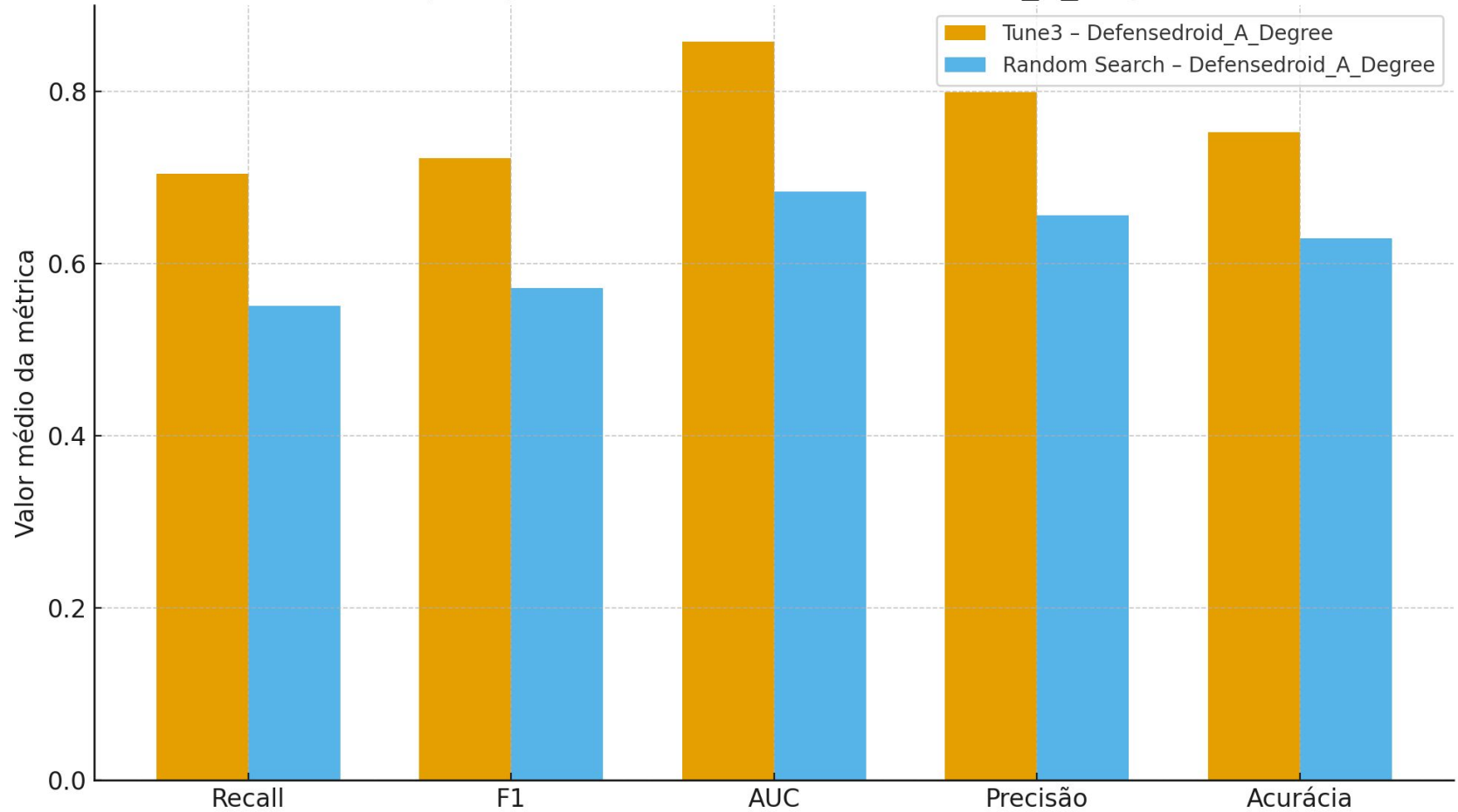
O Tune3 preservou ou melhorou o Recall enquanto reduziu o tempo acumulado de busca, indicando maior eficiência experimental.

Datasets avaliados
2

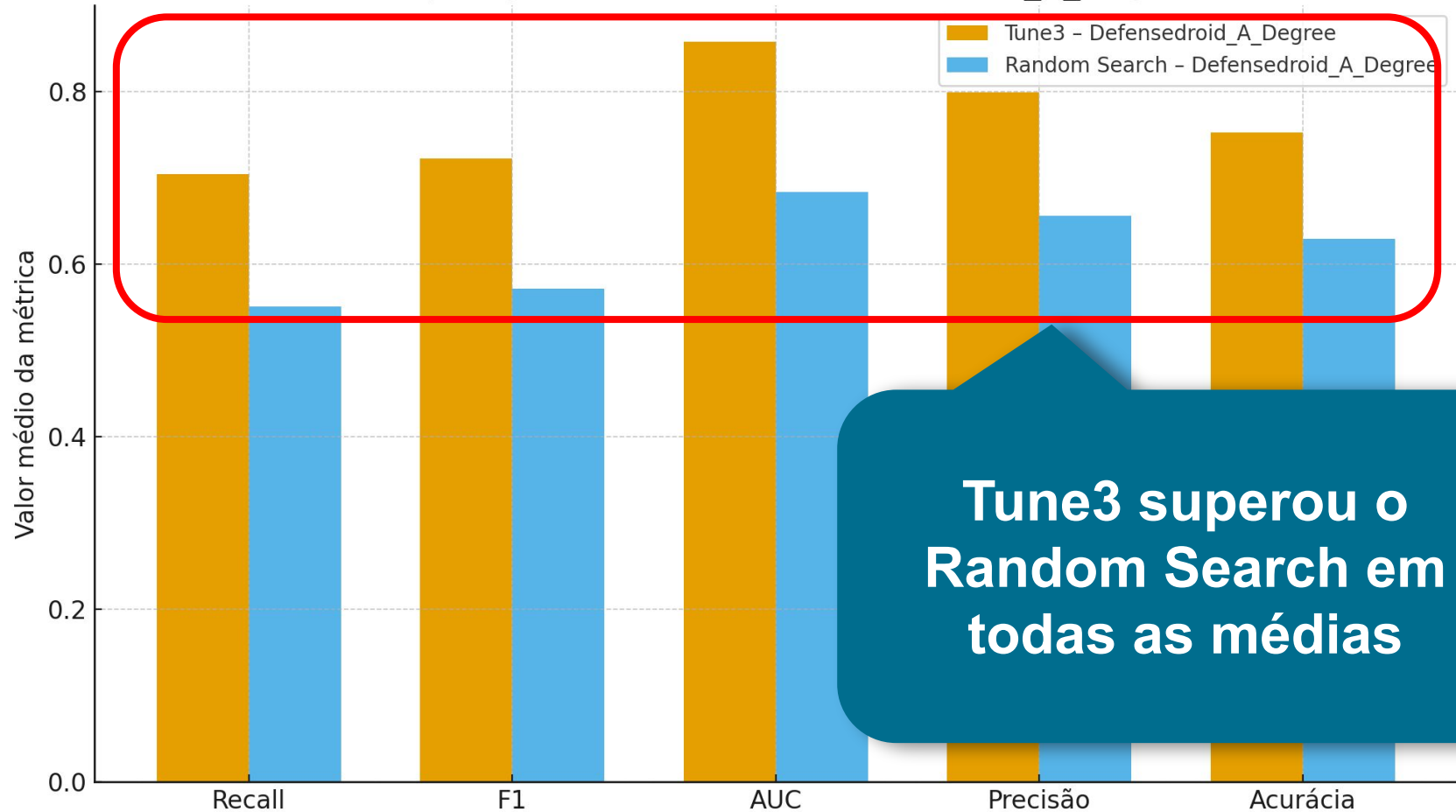
Ganho em Recall
+0,01 a +0,05

Tempo de busca
-34% a -44%

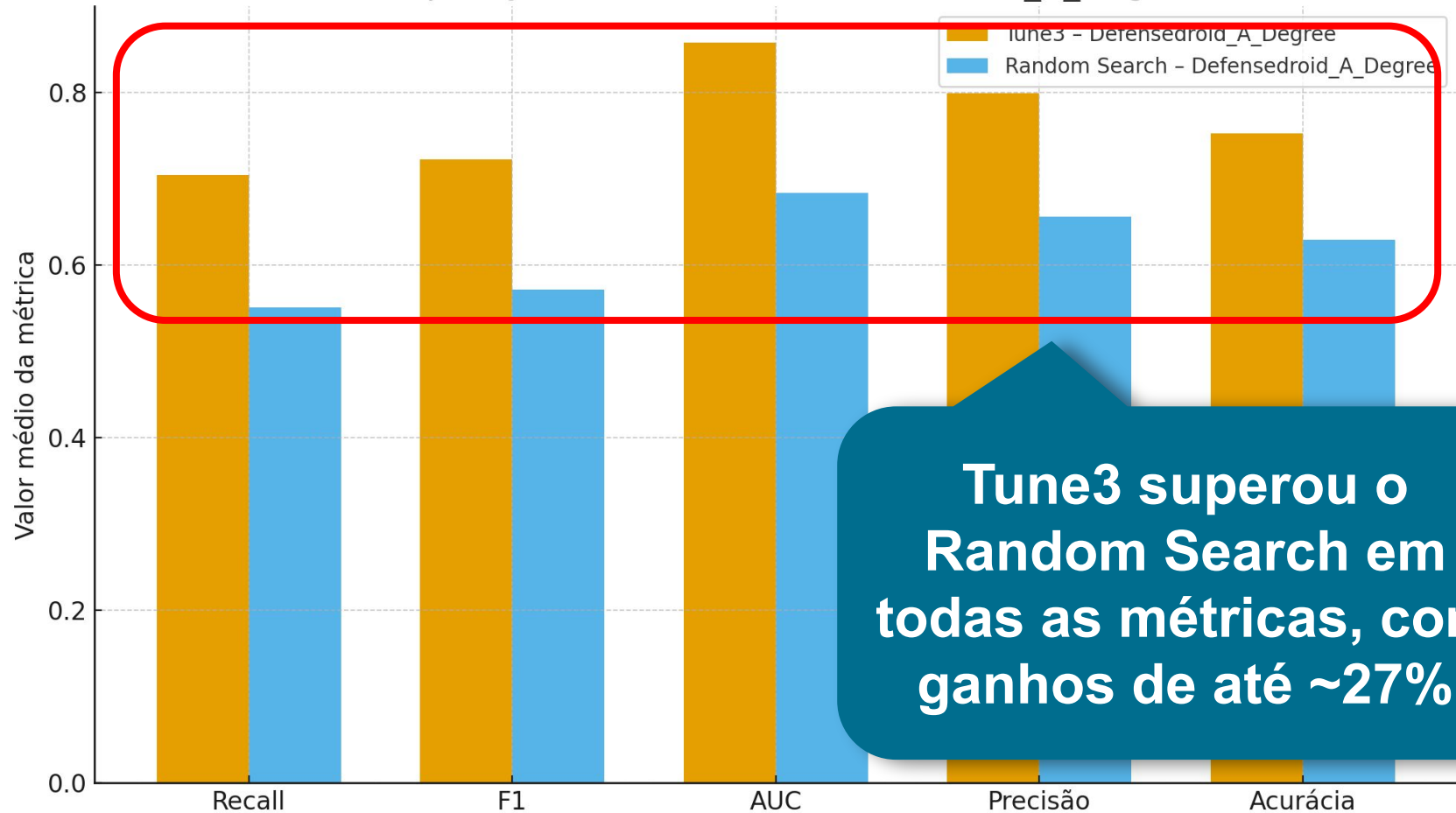
Comparação de métricas - Defensedroid_A_Degree



Comparação de métricas - Defensedroid_A_Degree



Comparação de métricas - Defensedroid_A_Degree



**Tune3 superou o
Random Search em
todas as métricas, com
ganhos de até ~27%**

Após reduzir o custo da busca por hiperparâmetros, o próximo desafio foi atacar outra fonte de custo: a alta dimensionalidade dos datasets.

Dataset

Dimensão (Samples × Features)

Exec (horas)

MH100k

~2,53 bilhões

28.583,6

KronoDroid Real

~22,3 mil

280,6

KronoDroid Emulator

~18,3 m

234,9

AndroCrawl

~13,6 m

182,7

Android Permissions

~4,0 milhões

73,8

DREBIN-215

~3,2 milhões

64,4

DefenseDroid PRS

~34,4 milhões

416,4

DD API Katz

~62,9 milhões

736,9

DD API Degree

~62,9 milhões

736,9

DD API Closeness

~44,8 milhões

532,8

Adroit

~1,9 milhões

49,4

**~ 3 anos de execução
na MalSynGen**

ARTIGO 3 / 4

Statistical Ranking: seleção de atributos por consenso (voto) na detecção de malware Android

Artigo completo · SBSEG 2026 · 11 datasets, validação por Wilcoxon

Métodos de Redução

- RFE (Recursive Feature Elimination)
- Semidroid
- Lasso



Métodos de Redução

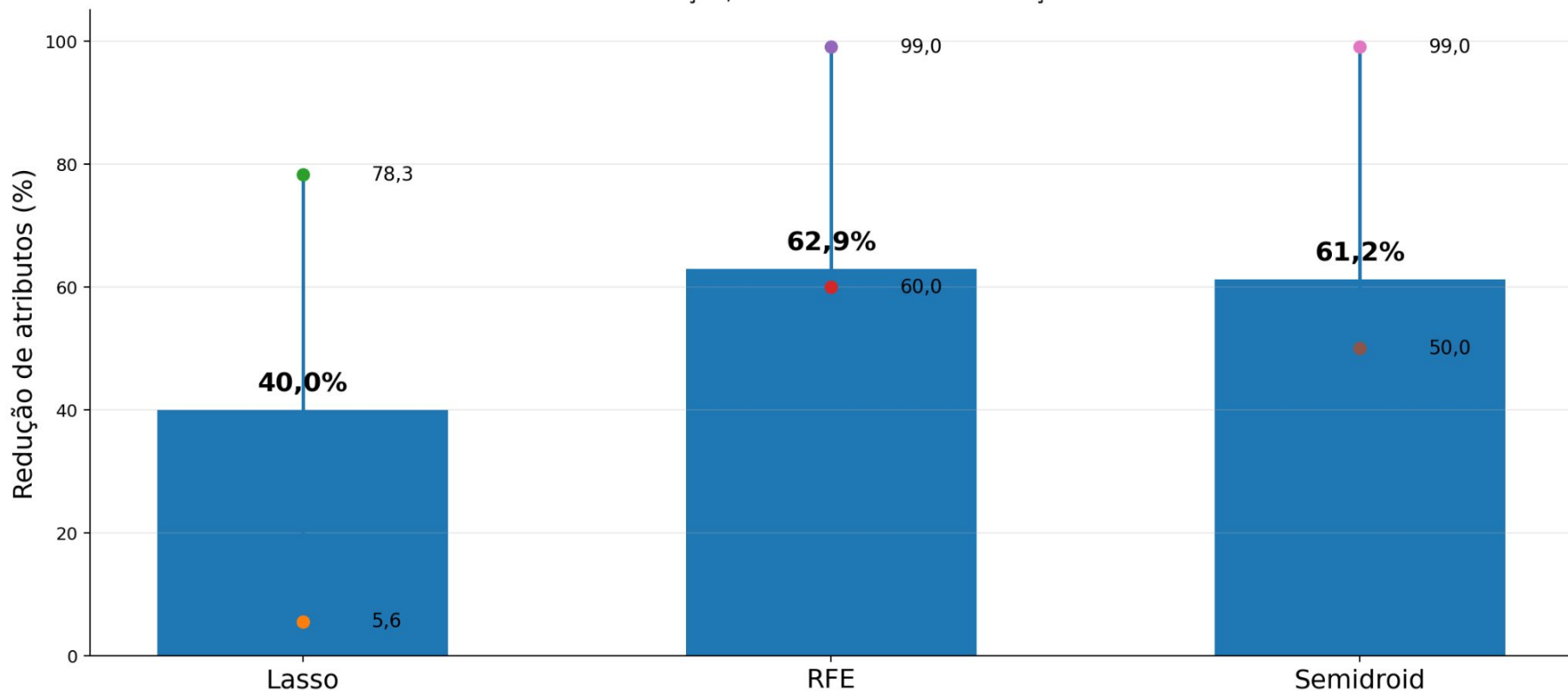
- RFE (Recursive Feature Elimination)
- Semidroid
- Lasso

**3 melhores métodos
MH-FSF**



Limitações dos métodos iniciais de redução

As barras indicam a mediana da redução; as linhas mostram a variação mínima-máxima observada.



Dataset	Year	Samples	Features	Benign	Malware
MH100k	2023	101.934	24.833	92.134	9.800
KronoDroid Real	2021	78.137	286	36.755	41.382
KronoDroid Emulator	2021	63.991	276	35.246	28.745
AndroCrawl	2013	96.744	141	86.574	10.170
Android Permissions	2018	26.864	151	9.077	17.787
DREBIN-215	2018	15.031	215	9.476	5.555
DefenseDroid PRS	2021	11.975	2.877	5.975	6.000
DD API Katz	2021	10.476	6.002	5.222	5.254
DD API Degree	2021	10.476	6.002	5.222	5.254
DD API Closeness	2021	10.476	4.274	5.222	5.254
Adroit	2017	11.476	166	8.058	3.418

Métodos de Redução: Statistical Ranking

Statistical Ranking

ETAPA 1: Pré-processamento

Remoção de baixa variância • Filtragem de correlação

ETAPA 2: Ranqueamento Estatístico

Chi-Quadrado • Informação Mútua • Random Forest

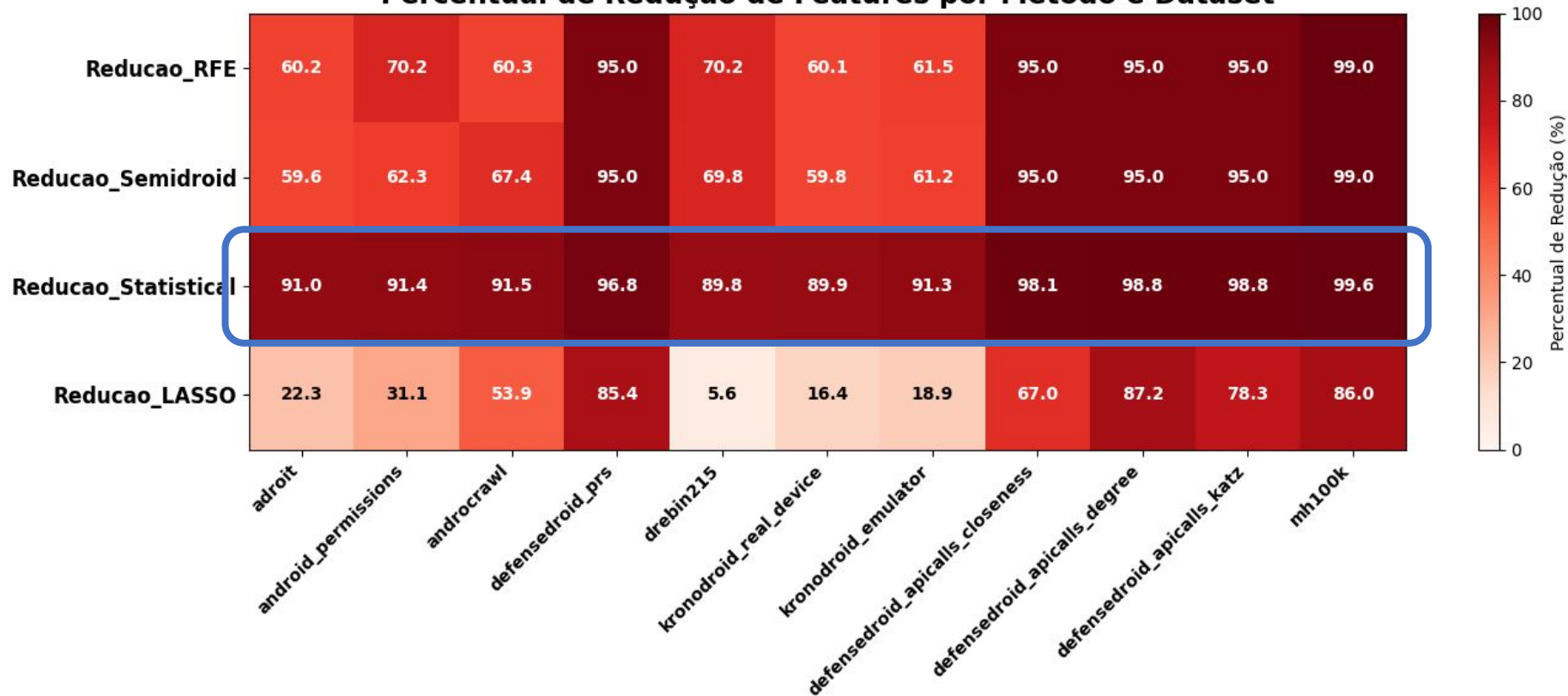
ETAPA 3: Votação por Método

Features com ≥ 2 votos (consenso robusto)

Dataset Reduzido

Dataset	Features Originais	RFE (feat %)	Semidroid (feat %)	Statistical (feat %)	Lasso (feat %)
adroit	166	66 (-60,2%)	67 (-59,6%)	15 (-91,0%)	129 (-22,3%)
android_permissions	151	45 (-70,2%)	46 (-62,3%)	13 (-91,4%)	104 (-31,1%)
androcrawl	141	56 (-60,3%)	57 (-67,4%)	12 (-91,5%)	65 (-53,9%)
defensedroid_prs	2877	143 (-95,0%)	144 (-95,0%)	92 (-96,8%)	421 (-85,4%)
drebin215	215	64 (-70,2%)	65 (-69,8%)	22 (-89,8%)	203 (-5,6%)
kronodroid_real_device	286	114 (-60,1%)	115 (-59,8%)	29 (-89,9%)	239 (-16,4%)
kronodroid_emulator	276	110 (-61,5%)	111 (-61,2%)	25 (-91,3%)	232 (-18,9%)
defensedroid_a_c	4274	213 (-95,0%)	214 (-95,0%)	81 (-98,1%)	1410 (-67,0%)
defensedroid_a_d	6002	300 (-95,0%)	301 (-95,0%)	75 (-98,8%)	771 (-87,2%)
defensedroid_a_k	6002	300 (-95,0%)	301 (-95,0%)	75 (-98,8%)	1305 (-78,3%)
PPGE-mh100k	24833	248 (-99,0%)	249 (-99,0%)	93 (-99,6%)	3477 (-86,0%)

Percentual de Redução de Features por Método e Dataset



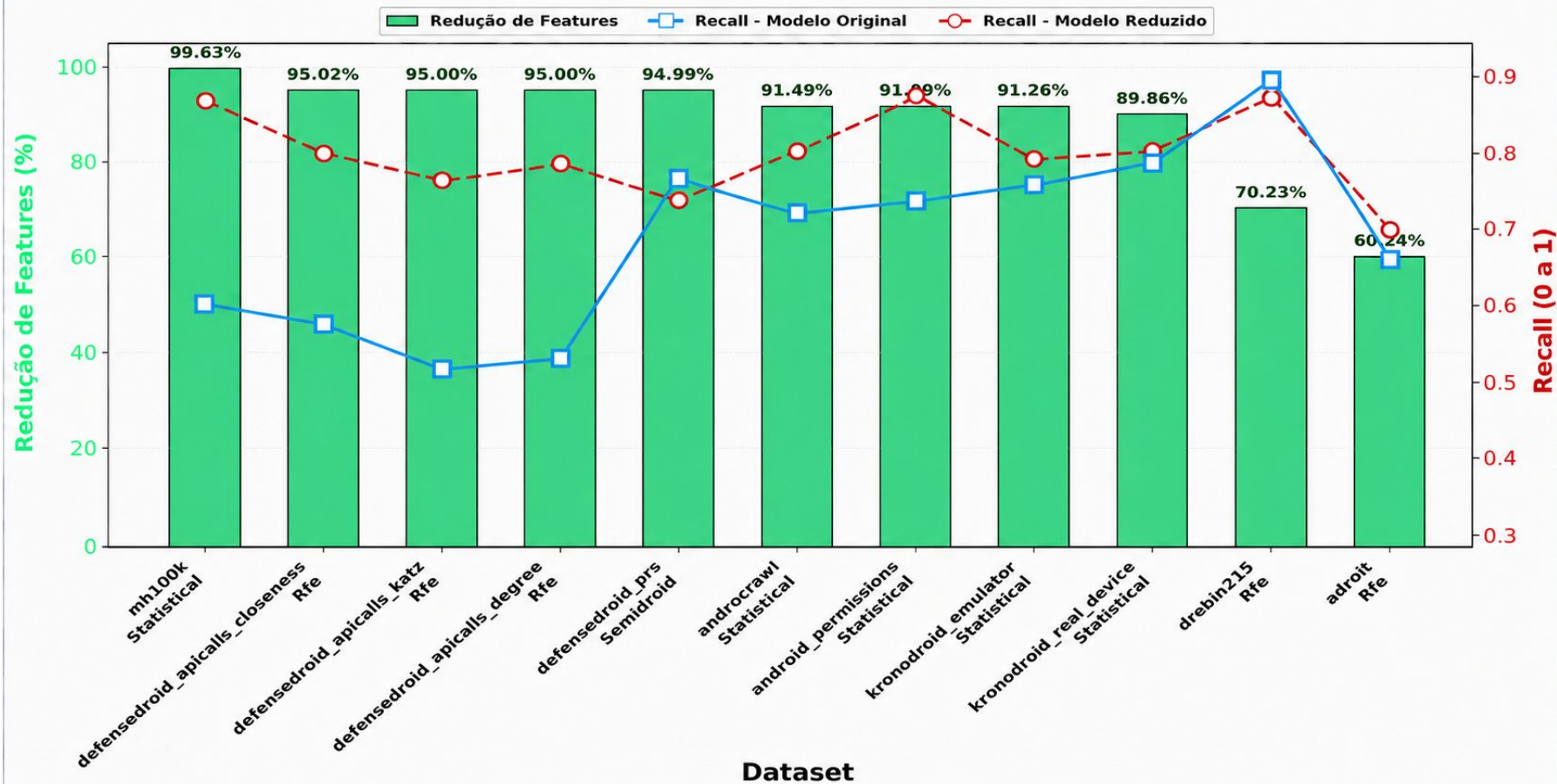
Avaliação das métricas

- Optamos por priorizar Recall, pois no contexto de malware, não detectar uma ameaça é mais crítico do que aumentar falsos positivos.

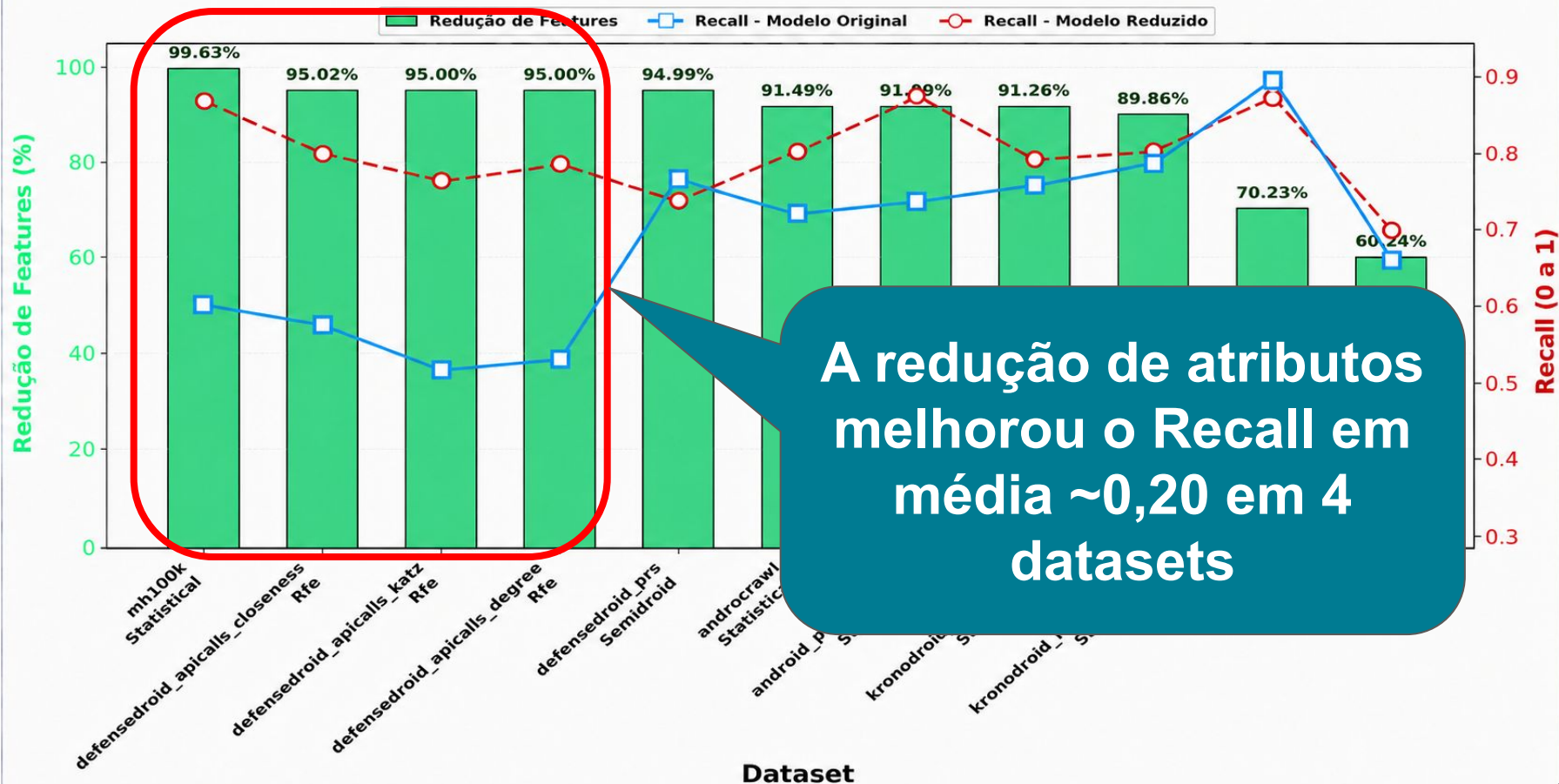
Avaliação das métricas

- Optamos por priorizar Recall, pois no contexto de malware, não detectar uma ameaça é mais crítico do que aumentar falsos positivos
- O F1-score foi utilizado como métrica de validação complementar para garantir que o ganho em Recall não ocorresse às custas de uma degradação extrema da Precisão.

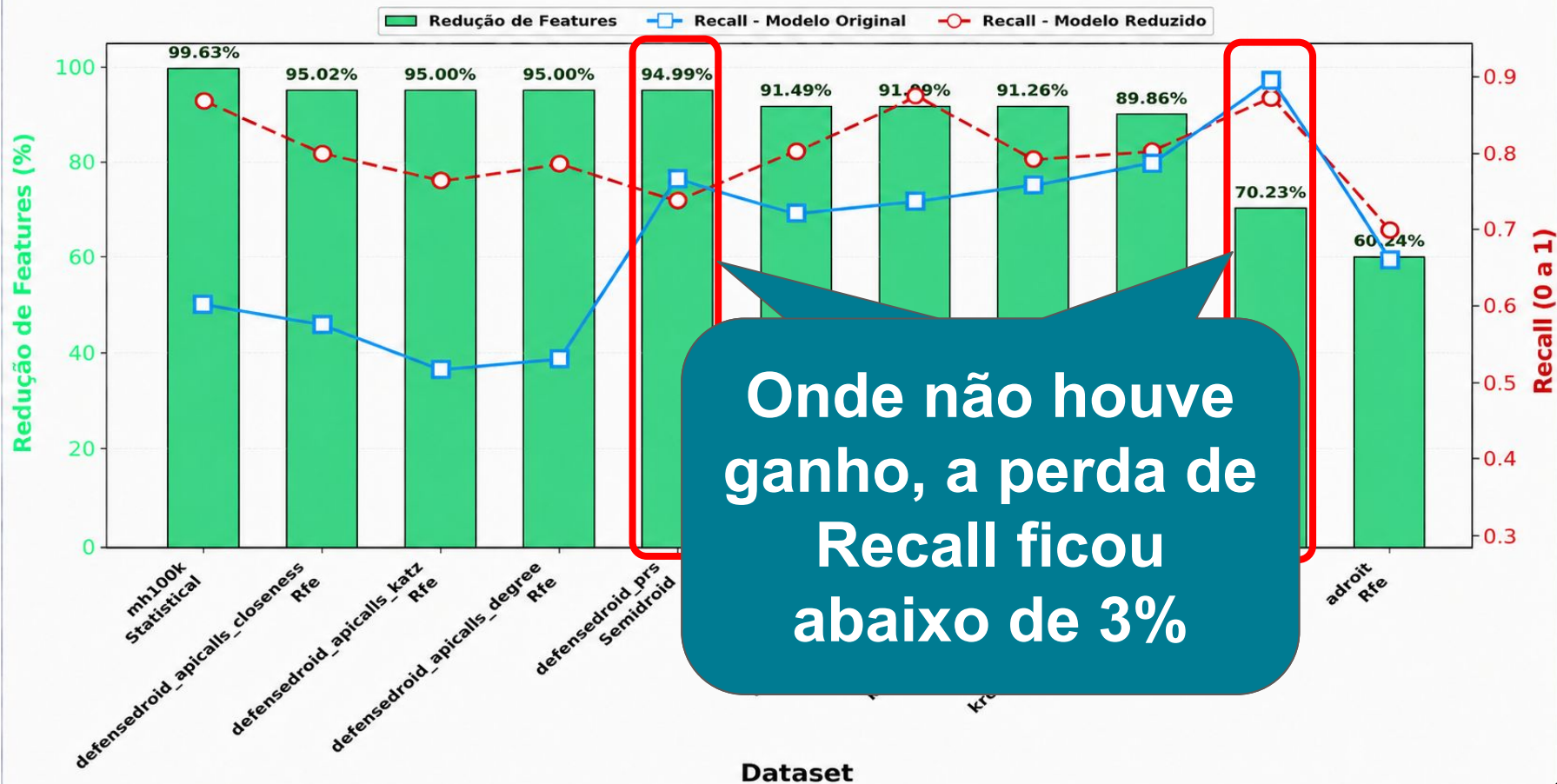
Impacto da Redução de Features no Recall Média entre classificadores por dataset



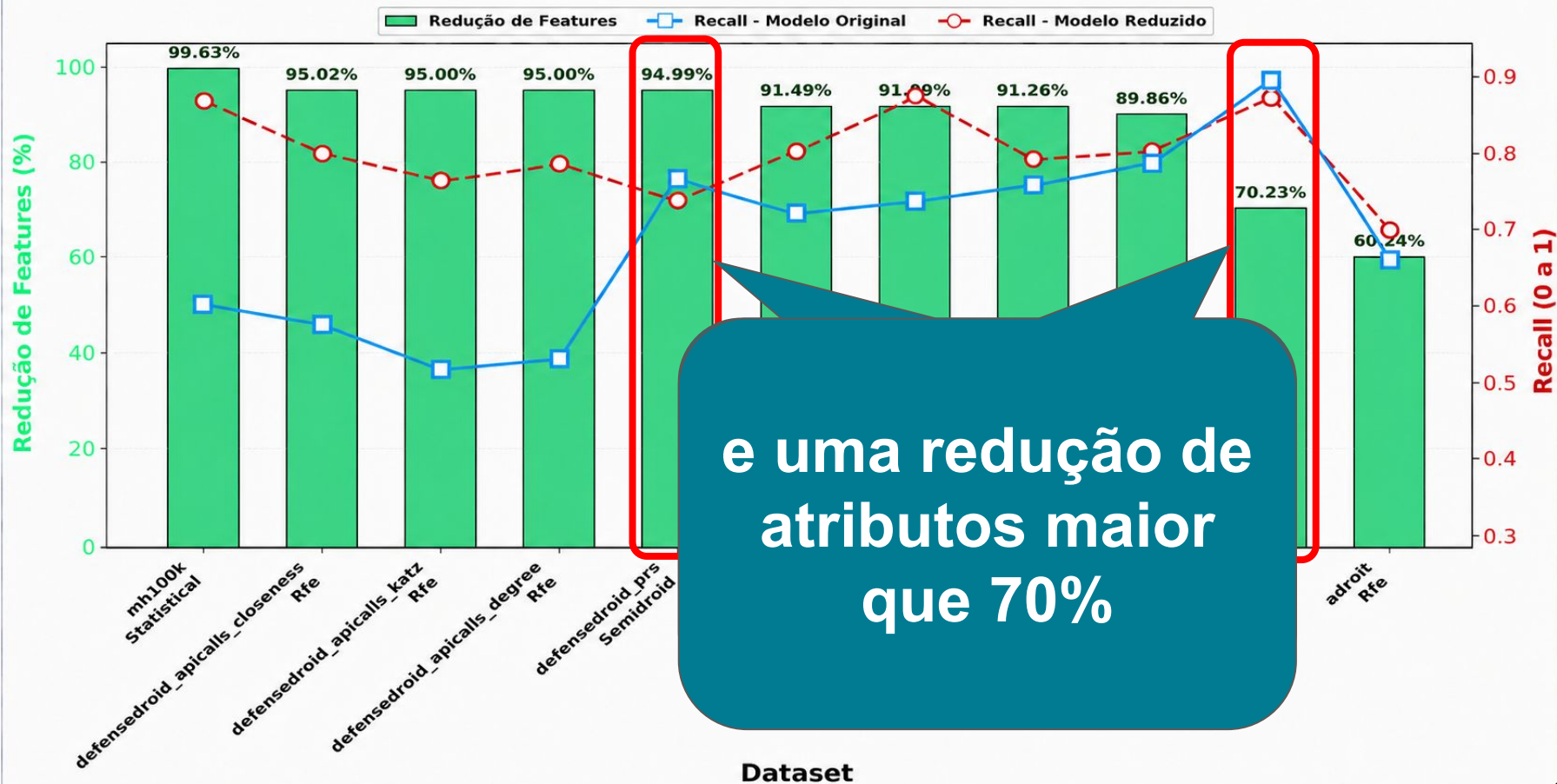
Impacto da Redução de Features no Recall Média entre classificadores por dataset



Impacto da Redução de Features no Recall Média entre classificadores por dataset



Impacto da Redução de Features no Recall Média entre classificadores por dataset



Resultados

Dataset	Método de redução
mh100k	statistical
kronodroid_real_device	statistical
kronodroid_emulator	statistical
drebin215	rfe
defensedroid_prs	semidroid
defensedroid_a_k	rfe
defensedroid_a_d	rfe
defensedroid_a_c	rfe
android_permissions	statistical
androcrawl	statistical
adroit	rfe

**Técnica
com
melhor
Recall
em 5 dos
11
datasets**

Resultados - Síntese

Menos dimensões, a mesma qualidade e uma fração do tempo

~80%

redução de tempo de
processamento

90–99,6%

Redução de
dimensionalidade

11

datasets reduzidos e
disponibilizados

Achado: remover atributos redundantes sem perda significativa de desempenho.

Com os datasets reduzidos definidos, o próximo passo foi ampliar as opções de datasets para a experimentação.

Como ampliar as opções de datasets para experimentação?

**Gerando versões balanceadas com
oversampling e undersampling, conforme
estratégias adotadas na literatura.**

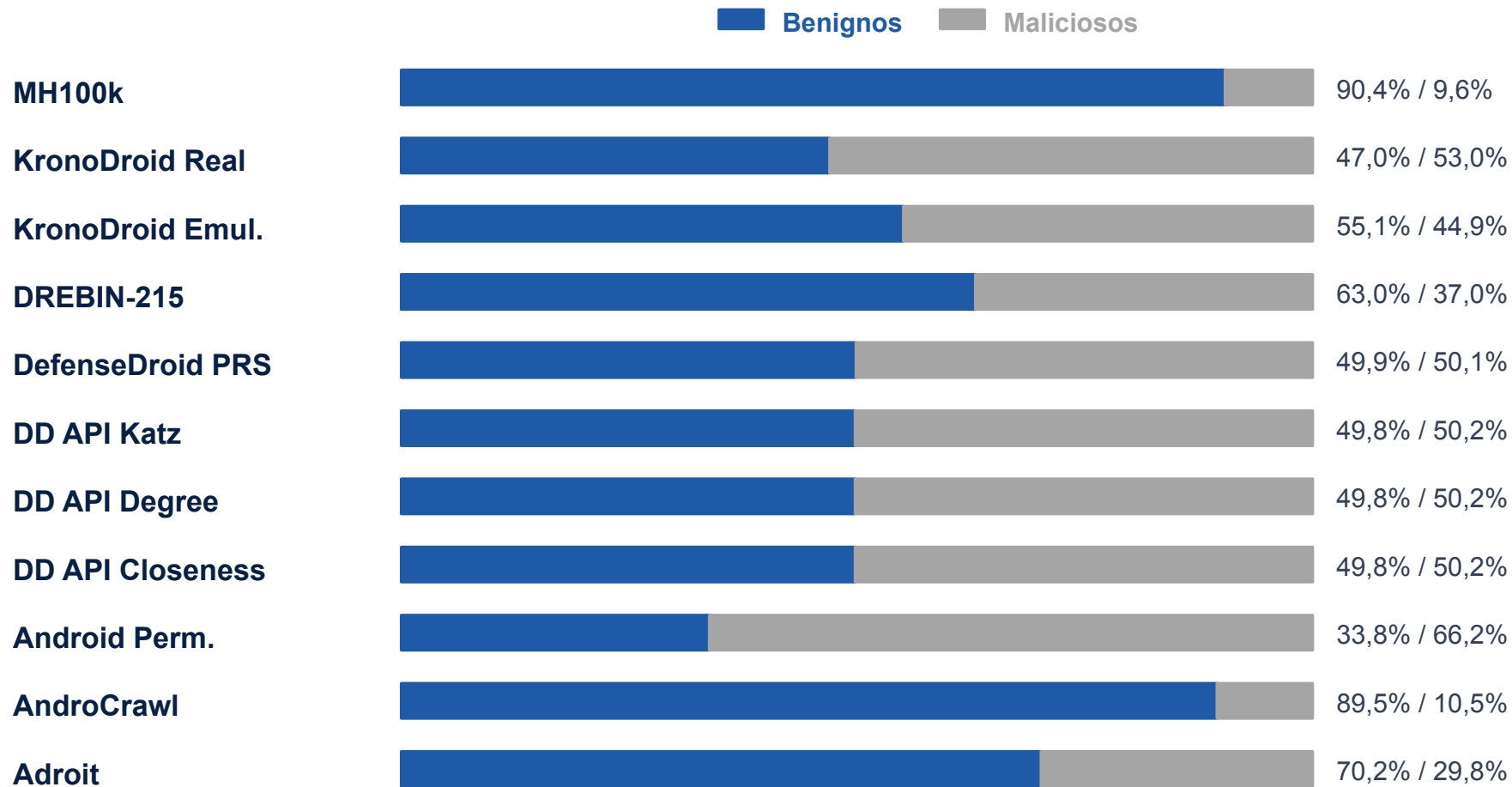
ARTIGO 4 / 4

When Balancing Harms: condições estruturais que degradam a detecção após o balanceamento

Artigo completo · SBSeg 2026 · 11 datasets, indicadores estruturais

Dataset	Features	(B/M)
mh100k	93	92,134 / 9,800 (9,4:1)
kronodroid_real_device	29	36,755 / 41,382 (1:1.1)
kronodroid_emulator	25	35,246 / 28,745 (1.2:1)
drebin215	64	9,476 / 5,555 (1.7:1)
defensedroid_prs	144	5,975 / 6,000 (~1:1)
defensedroid_apicalls_katz	300	5,222 / 5,254 (~1:1)
defensedroid_apicalls_degree	300	5,222 / 5,254 (~1:1)
defensedroid_apicalls_closeness	213	5,222 / 5,254 (~1:1)
android_permissions	13	9,077 / 17,787 (1:2)
androcrawl	12	86,574 / 10,170 (8,5:1)
adroit	66	8,058 / 3,418 (2,4:1)

Dataset	Features	(B/M)
mh100k	93	92,134 / 9,800 (9,4:1)
kronodroid_real_device	29	36,755 / 41,382 (1:1.1)
kronodroid_emulator	25	35,246 / 28,745 (1.2:1)
drebin215	64	9,476 / 5,555 (1.7:1)
defensedroid_prs	144	5,975 / 6,000 (~1:1)
defensedroid_apicalls_katz	300	5,222 / 5,254 (~1:1)
defensedroid_apicalls_degree	300	5,222 / 5,254 (~1:1)
defensedroid_apicalls_closeness	213	5,222 / 5,254 (~1:1)
android_permissions	13	9,077 / 17,787 (1:2)
androcrawl	12	86,574 / 10,170 (8,5:1)
adroit	66	8,058 / 3,418 (2,4:1)



Técnicas de Balanceamento

Mesmo pipeline em três versões: original · reduzido · balanceado

Undersampling

RandomUnder · NearMiss · ENN · Tomek Links ·
ENN + RandomUnder · Tomek + RandomUnder

Oversampling

RandomOver · SMOTE · Borderline-SMOTE ·
ADASYN

Recall como métrica primária (falsos negativos = ameaças não detectadas); $F1 \geq 0,70$ como critério mínimo

Técnicas de Balanceamento

Mesmo pipeline em três versões: original · reduzido · balanceado

Undersampling

RandomUnder · NearMiss · ENN · Tomek Links ·
ENN + RandomUnder · Tomek + RandomUnder

Oversampling

RandomOver · SMOTE · Borderline-SMOTE ·
ADASYN

Recall como métrica primária (falsos negativos = ameaças não detectadas); $F1 \geq 0,70$ como critério mínimo

Escala dos Experimentos

- 130 variações de datasets
- ~230 execuções de experimentos
- ~1 mês de execução contínua



Dataset	Undersampling (B/M + Técnica)	Oversampling (B/M + Técnica)
mh100k	9,800 / 9,800 (1:1, ENN + RandomUnder)	92,134 / 92,134 (1:1, RandomOver)
kronodroid_real_device	36,755 / 36,755 (1:1, ENN + RandomUnder)	40,873 / 41,382 (~1:1, ADASYN)
kronodroid_emulator	28,745 / 28,745 (1:1, ENN + RandomUnder)	35,246 / 35,246 (1:1, SMOTE)
drebin215	5,555 / 5,555 (1:1, ENN + RandomUnder)	9,476 / 9,476 (1:1, SMOTE)
defensedroid_prs	4,827 / 4,827 (1:1, ENN + RandomUnder)	6,000 / 6,000 (1:1, RandomOver)
defensedroid_a_k	4,571 / 4,571 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
defensedroid_a_d	4,571 / 4,571 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
defensedroid_a_c	4,563 / 4,563 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
android_permissions	7,379 / 7,379 (1:1, ENN + RandomUnder)	17,787 / 17,787 (1:1, SMOTE)
androcrawl	10,170 / 10,170 (1:1, NearMiss)	86,574 / 86,574 (1:1, SMOTE)
adroit	3,418 / 3,418 (1:1, NearMiss)	8,058 / 8,058 (1:1, RandomOver)

Dataset	Undersampling (B/M + Técnica)	Oversampling (B/M + Técnica)
mh100k	9,800 / 9,800 (1:1, ENN + RandomUnder)	92,134 / 92,134 (1:1, RandomOver)
kronodroid_real_device	36,755 / 36,755 (1:1, ENN + RandomUnder)	40,873 / 41,382 (~1:1, ADASYN)
kronodroid_emulator	28,745 / 28,745 (1:1, ENN + RandomUnder)	35,246 / 35,246 (1:1, SMOTE)
drebin215	5,555 / 5,555 (1:1, ENN + RandomUnder)	9,476 / 9,476 (1:1, SMOTE)
defensedroid_prs	4,827 / 4,827 (1:1, ENN + RandomUnder)	6,000 / 6,000 (1:1, RandomOver)
defensedroid_a_k	4,571 / 4,571 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
defensedroid_a_d	4,571 / 4,571 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
	4,563 / 4,563 (1:1, ENN + RandomUnder)	5,254 / 5,254 (1:1, RandomOver)
	,379 / 7,379 (1:1, ENN + RandomUnder)	17,787 / 17,787 (1:1, SMOTE)
	0,170 / 10,170 (1:1, NearMiss)	86,574 / 86,574 (1:1, SMOTE)
	3,418 / 3,418 (1:1, NearMiss)	8,058 / 8,058 (1:1, RandomOver)

**22 datasets
balanceados**

Resultados

O efeito do balanceamento depende do dataset

18 / 22

datasets com
manteve/melhorou
as métricas

0,87 → 0,99

Recall no
MH100k

~70%

redução de
tempo

Resultados

Dataset reduzido + Undersampling reduziu o custo computacional significativamente

~~28.583 h~~ → **130 h** ≈ 220× mais rápido

MH100k — maior dataset, maior ganho

O mesmo padrão se repete nos 11 datasets: de centenas de horas para dezenas.

apicalls_degree

~~736 h~~ → **63 h**

apicalls_closeness

~~532 h~~ → **53 h**

defensedroid_prs

~~416 h~~ → **47 h**

kronodroid_real

~~280 h~~ → **54 h**

Contribuição do experimento para a comunidade

33 datasets gerados para experimentações

33

datasets gerados e
disponibilizados

11 reduzidos

Dimensionalidade
cortada com
qualidade preservada

22 balanceados

Prontos para os
experimentos.



Quatro artigos, uma trajetória de pesquisa

A apresentação segue a ordem cronológica dos trabalhos



Best Paper ERRC/WRSeg



O trabalho intitulado “**Otimização de hiperparâmetros da DroidAugmentor para geração de dados sintéticos de malware Android**”, de autoria de **Angelo Gaspar Diniz Nogueira, Lucas Ferreira Areais de Oliveira, Anna Luiza Gomes da Silva, Diego Kreutz, Rodrigo Brandão Mansilha (Unipampa)**, recebeu o prêmio de **Melhor Artigo WRSeg**, apresentado no VIII Workshop Regional de Segurança da Informação e de Sistemas Computacionais (WRSeg 2024) da XXI Escola Regional de Redes de Computadores (ERRC 2024), realizado de 27 a 29 de novembro de 2024.

Rio Grande, 29 de novembro de 2024

Coordenação Geral

Prof. Dr. Bruno Lopes Dalmazo (FURG)
Prof. Dr. Pedro de Botelho Marcos (FURG)
Prof. Dr. Vinicius Garcia Pinto (FURG)

Comitê de Programa do WRSeg

Prof. Dr. Bruno Lopes Dalmazo (FURG)
Prof. Dr. Charles Christian Miers (UDESC)



FURG
UNIVERSIDADE FEDERAL
DO RIO GRANDE



UDESC



MENÇÃO HONROSA

Certificamos que o trabalho **“Tune3: Otimização Multiestágio de Hiperparâmetros com Redução de Custo e Maior Estabilidade em Modelos de Detecção de Malware Android”** de autoria de Lucas Ferreira, Angelo Nogueira, Diego Kreutz, Dionatan Schmidt e Rodrigo Mansilha foi o **3º colocado (Menção Honrosa)** da categoria de **Melhor Artigo** do 9º Workshop Regional de Segurança da Informação e de Sistemas Computacionais (WRSeg 2025)

Prof. Dr. Luís Augusto Dias Knob
Coordenador do Comitê de Programa

Prof. Dr. Muriel Figueredo Franco
Representante da Comissão de Avaliação

Organização

Unisenac
Centro Universitário RS

senac

Fecomércio
IFEP
Sesc
Sindicatos Empresariais

Trabalhos futuros

- Avaliar o Tune3 com os 33 datasets
- Comparar o Tune3 com Optuna e Hyperopt
- Formalizar matematicamente o Tune3

CONCLUSÃO

Escolher melhor vale mais que computar mais.

Diagnóstico, busca progressiva, redução de dimensionalidade e datasets reprodutíveis transformam experimentos custosos em experimentos viáveis.

Obrigado.

Perguntas e discussão

Lucas Ferreira · PPGES — UNIPAMPA

tune3-hpo.github.io

