MalSynGen: redes neurais artificiais na geração de dados tabulares sintéticos para detecção de malware

Angelo Gaspar, Diego Kreutz
Hendrio Bagança
Rodrigo Mansilha
Kayuã Oleques Paim

PPGES
Programa de Pós-Graduação em Engenharia de Software

LEA
Laboratório de Estudos Avançados em Computação

unipampa
Universidade Federal do Pampa

UFAM

UFRGS
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Motivação

80% dos projetos de IA falham por questões que envolvem dados

Top 10 Reasons Why AI Projects Fail

Motivação

Dataset	Ano	Características	Amostras		
			M	B	P
Drebin-215	2012/2018	215	5555	9476	0.5862
AndroCrawl	2013	86	10170	86562	0.1175
DefenseDroid	2021	2938	6000	5975	1.00033
MH-100K	2023	24833	9800	92134	0.1063

Motivação

Dataset	Ano	Car	Amostras		P
Drebin-215	2012/2018				0.5862
AndroCrawl	2013	86	10170	86562	0.1175
DefenseDroid	2021	2938	6000	5975	1.00033
MH-100K	2023	24833	9800	92134	0.1063

Defasado em relação ao malware atual

Motivação

Dataset	Ano	Características	Amostras		
Drebin-215	2012/2018	215			
AndroCrawl	2013	86	10170	86562	0.1175
DefenseDroid	2021	2938	6000	5975	1.00033
MH-100K	2023	24833	9800	92134	0.1063

Número limitado de características

Motivação

Dataset	Ano	Características	Amostras		
			M	B	P
Drebin-215	2012/2018	215	5555	9476	0.5862
AndroCrawl	2013	86			
DefenseDroid	2021	2938			
MH-100K	2023	24833	9800	92134	0.1063

Diversas vezes mais características

Motivação

Dat		rísticas	Amostras		
			M	B	P
Dreb		5	5555	9476	0.5862
Andro		6	10170	86562	0.1175
DefenseDroid	2021	2938	6000	5975	1.00033
MH-100K	2023	24833	9800	92134	0.1063

Desbalanceamento entre malware e benignos

Motivação

Dat		rísticas	Amostras		
			M	B	P
Dreb		5	5555	9476	0.5862
Andro		6	10170	86562	0.1175
DefenseDroid	2021	2938	6000	5975	1.00033
MH-100K	2023	24833	9800	92134	0.1063

Desbalanceamento entre malware e benignos

Motivação


VIRUSTOTAL
Analyse suspicious files, domains, IPs and URLs to detect malware and other breaches, automatically share them with the security community.
FILE
URL
SEARCH
Choose file

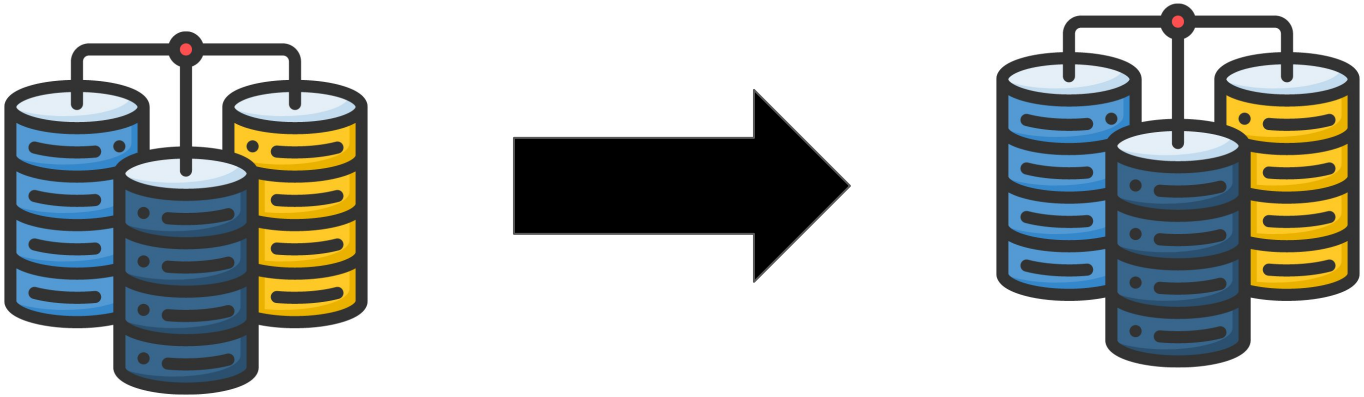
Motivação


VIRUSTOTAL
Analyse suspicious files, domains, IPs and URLs to detect malware and other
breaches, automatically share them with the s urity community.
FILE
URL
SEARCH
Serviço do VirusTotal permite apenas 250 rotulações por dia

Motivação


VIRUSTOT
Analyse suspicious files, domains, IPs and URLs to detect malware an
breaches, automatically share them with the security community
FILE
URL
SEARCH
Choose file
200 dias para rotular 50000 amostras

Aumento de dados



datasets reais

datasets sintéticos

Aumento de dados

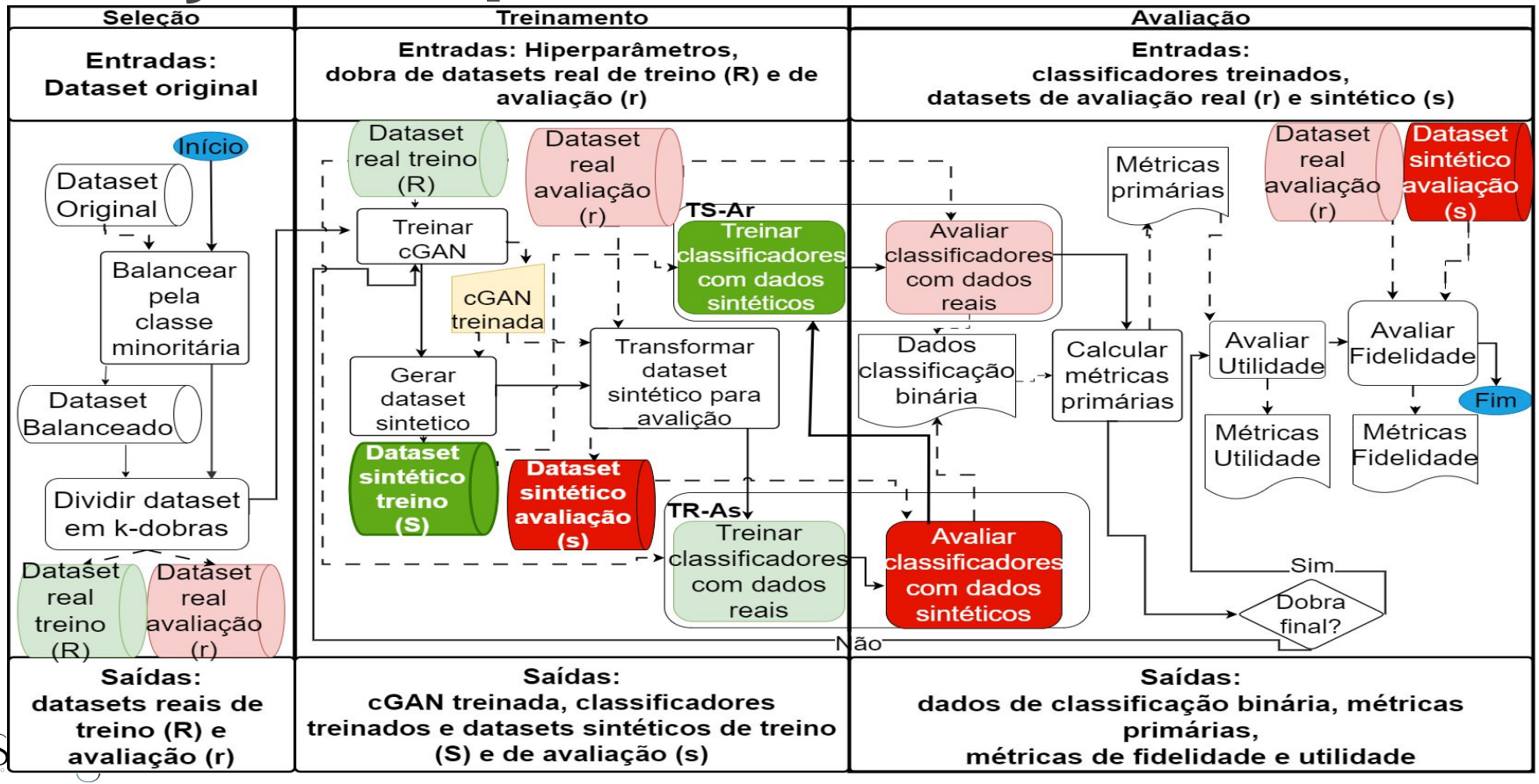
datasets reais

datasets sintéticos

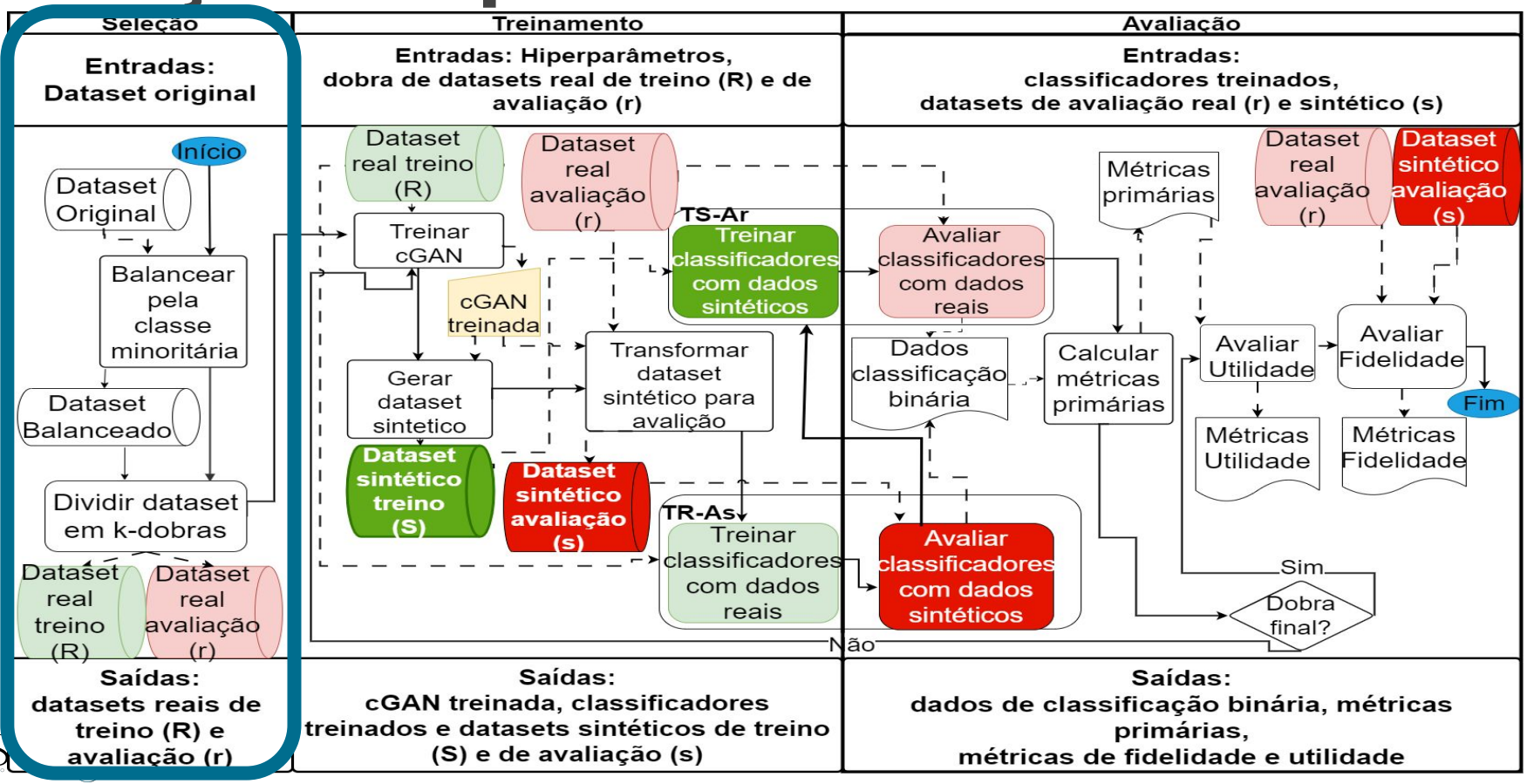
Aumento de dados: benefícios

- Desempenho aprimorado do modelo
- Evitar *overfitting*
- Maior privacidade dos dados

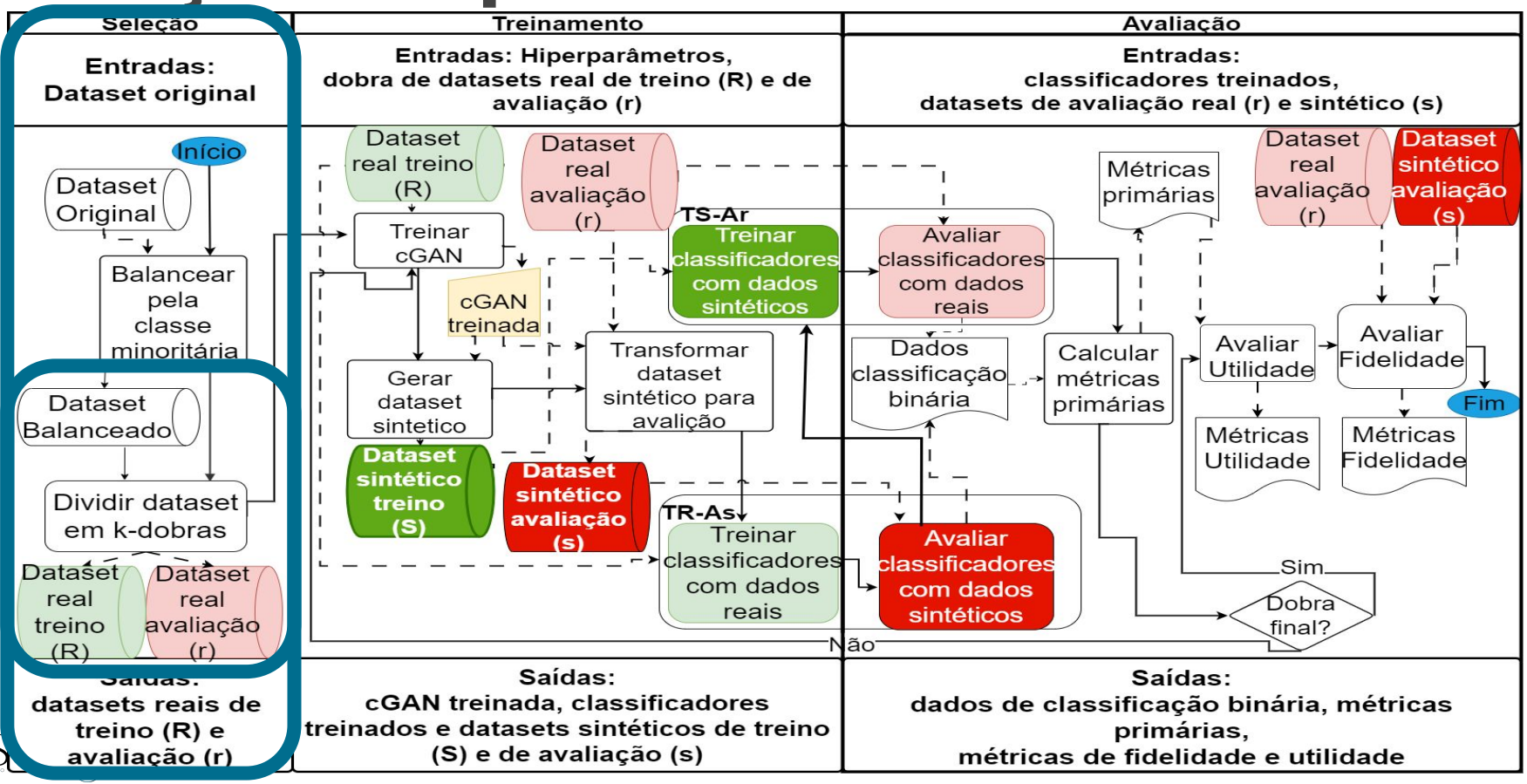
Solução Proposta



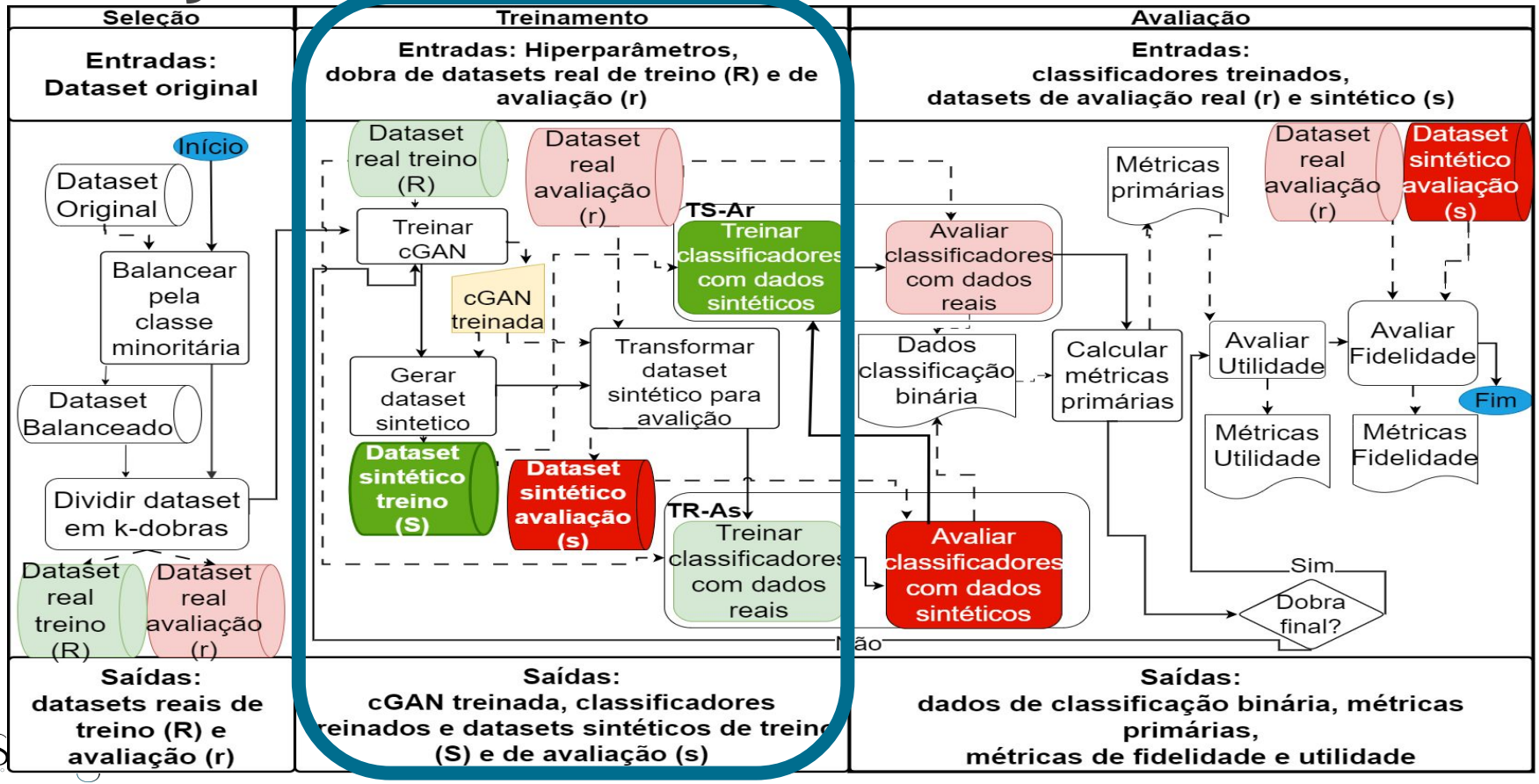
Solução Proposta



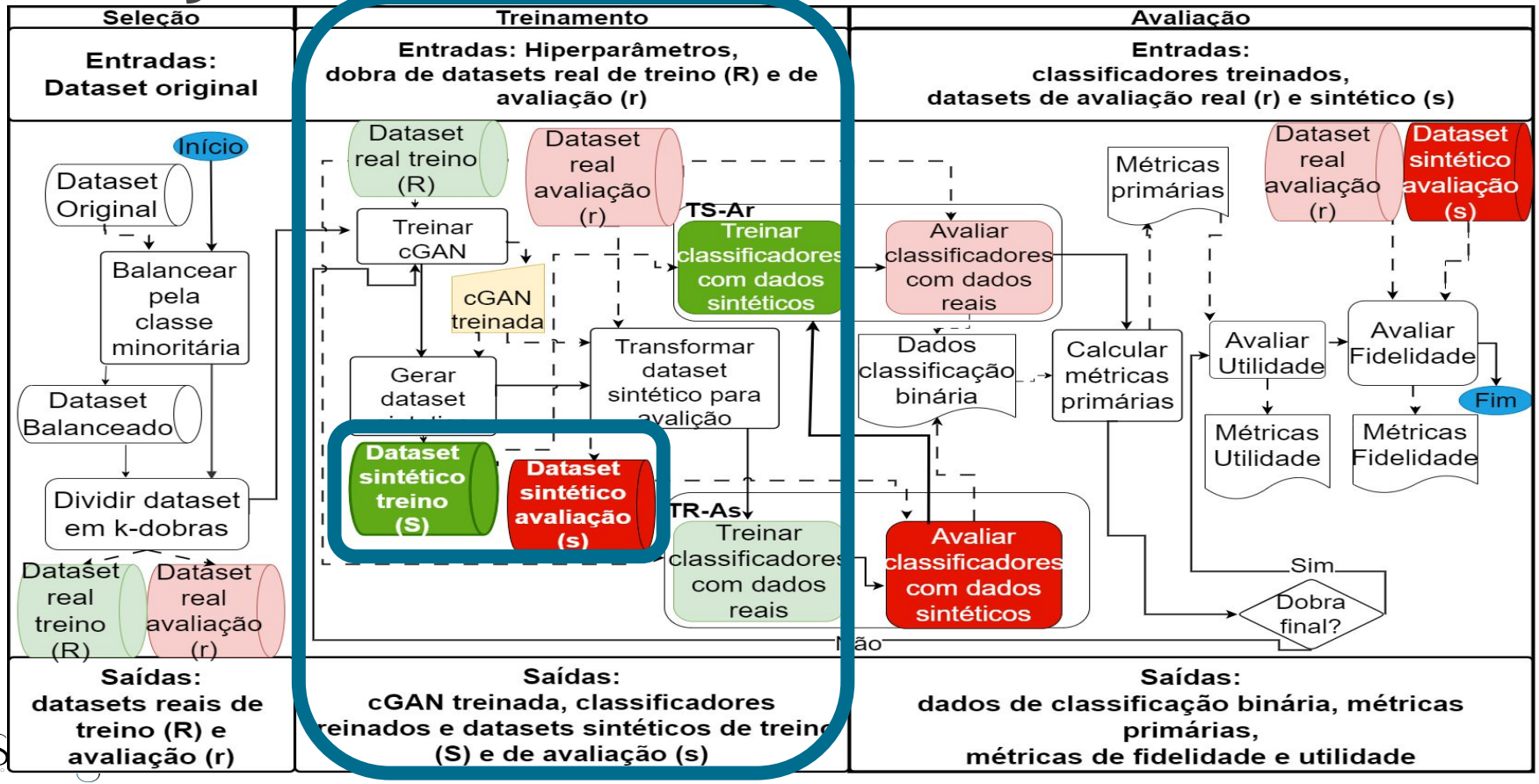
Solução Proposta



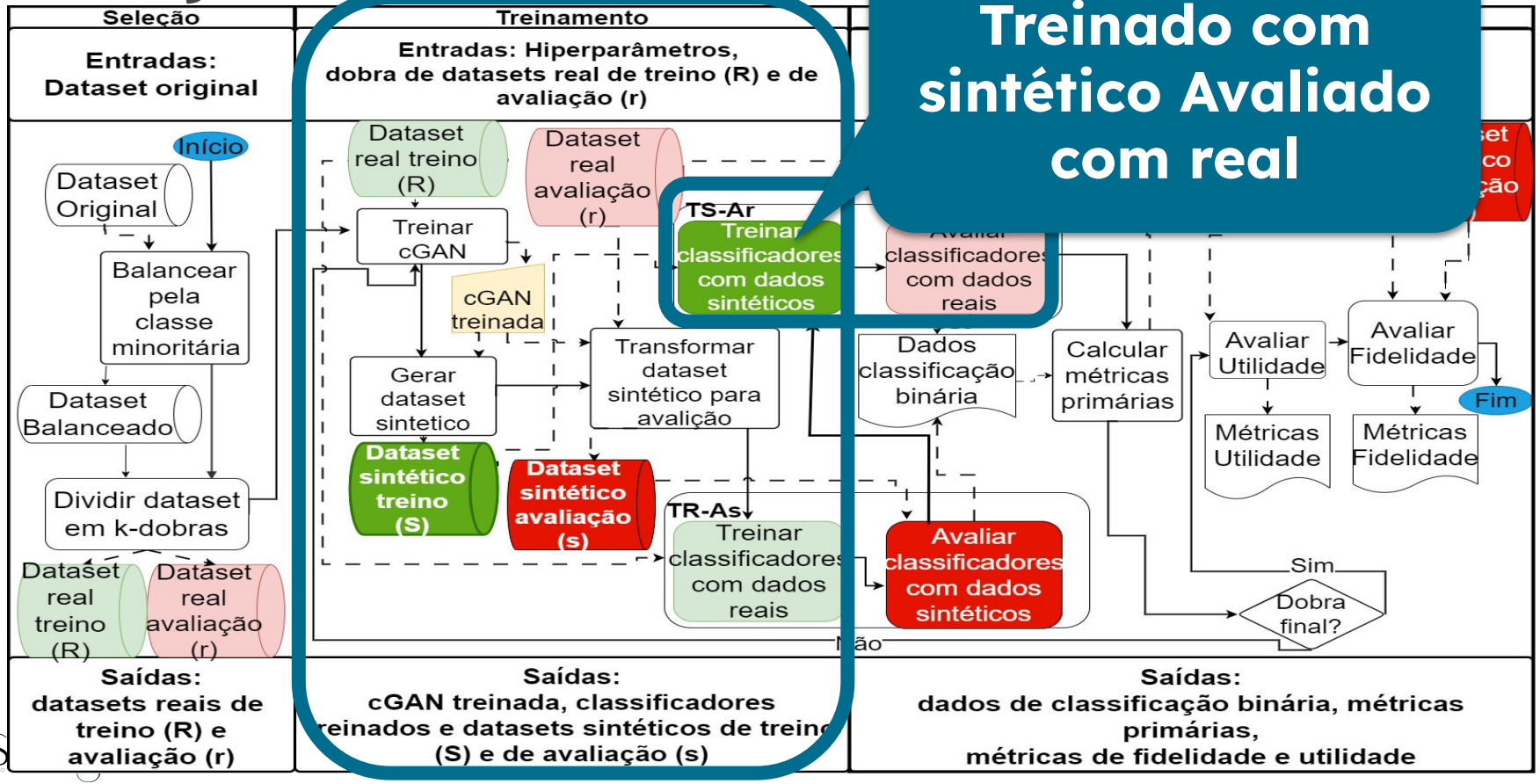
Solução Proposta



Solução Proposta

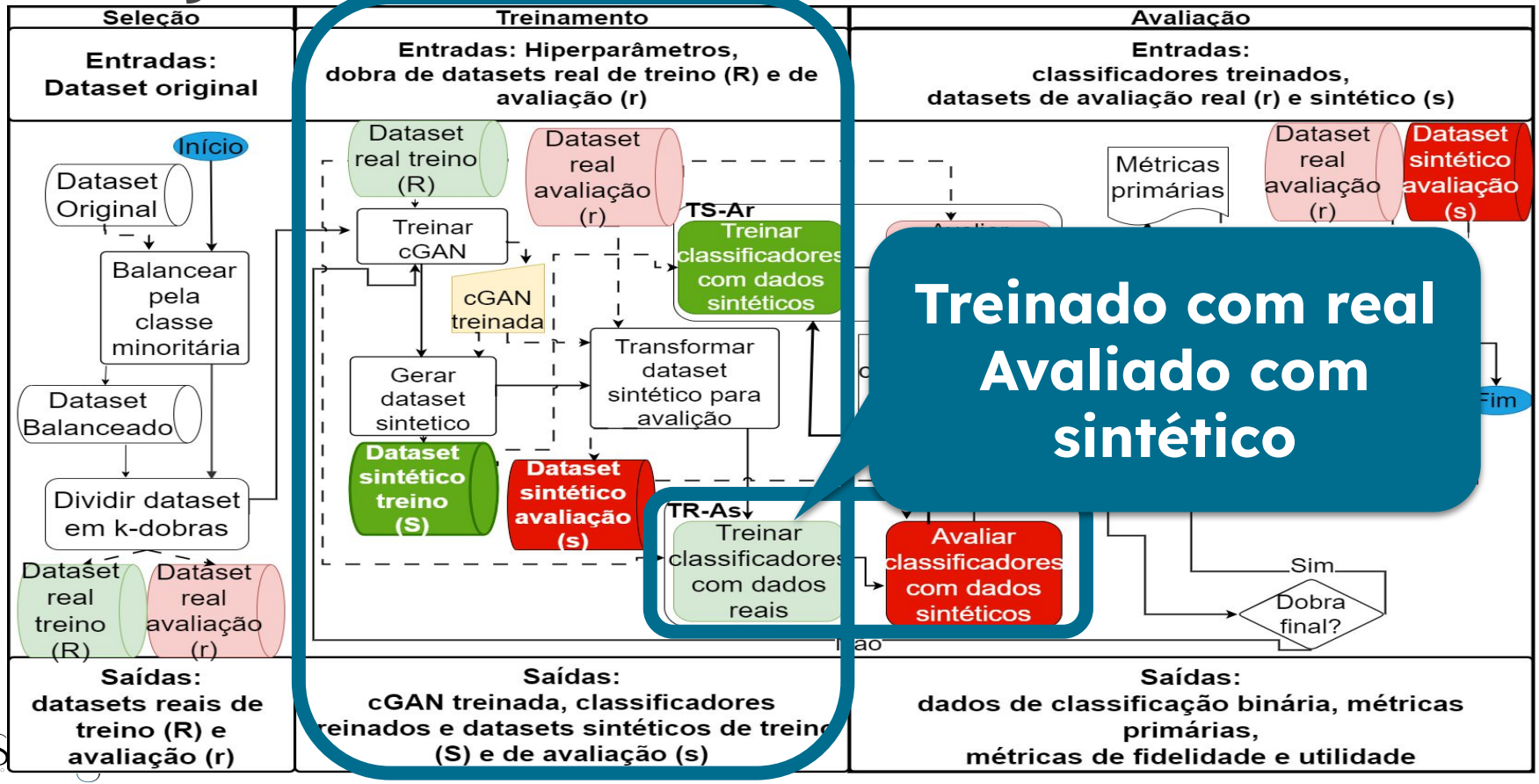


Solução Proposta



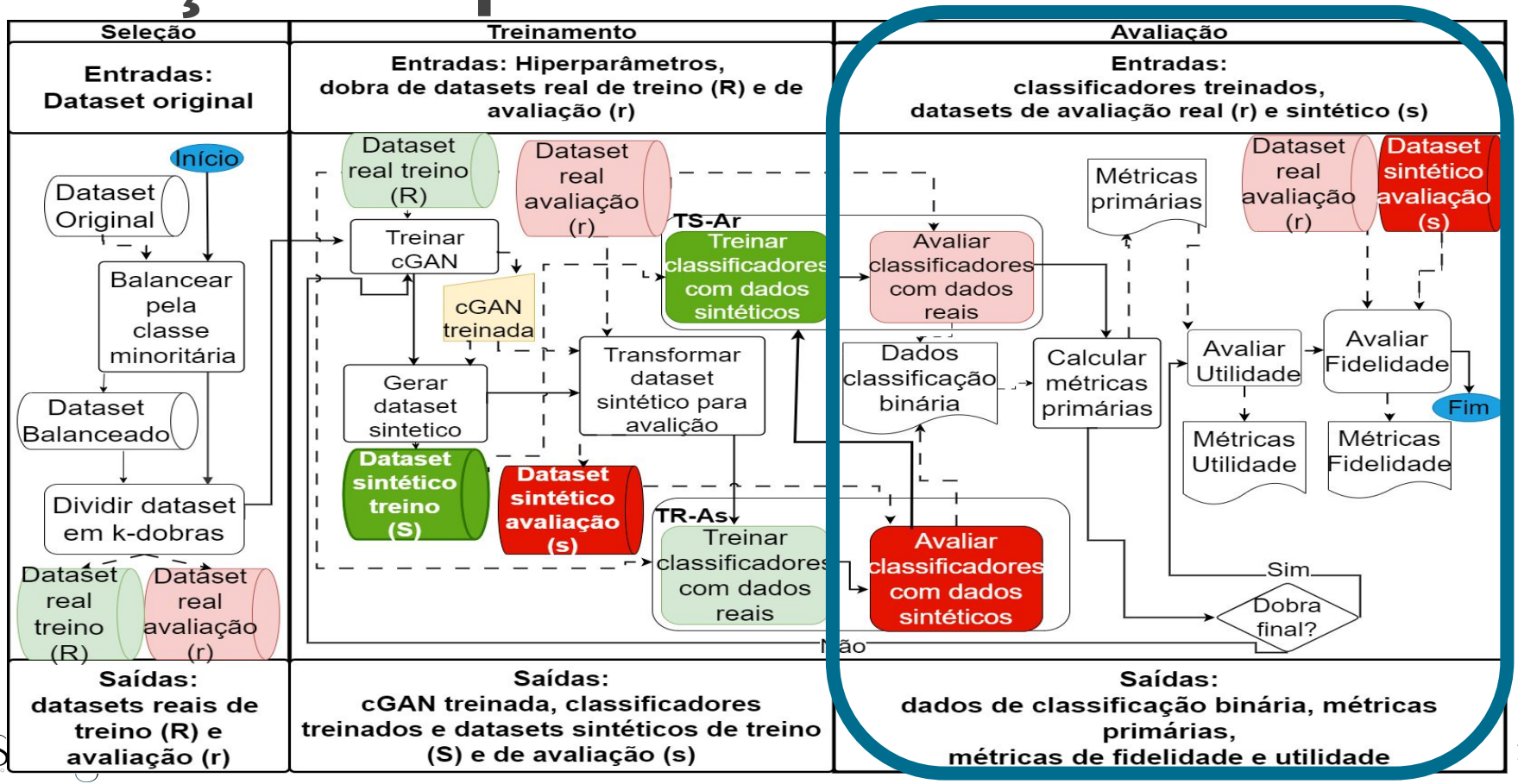
Treinado com sintético Avaliado com real

Solução Proposta

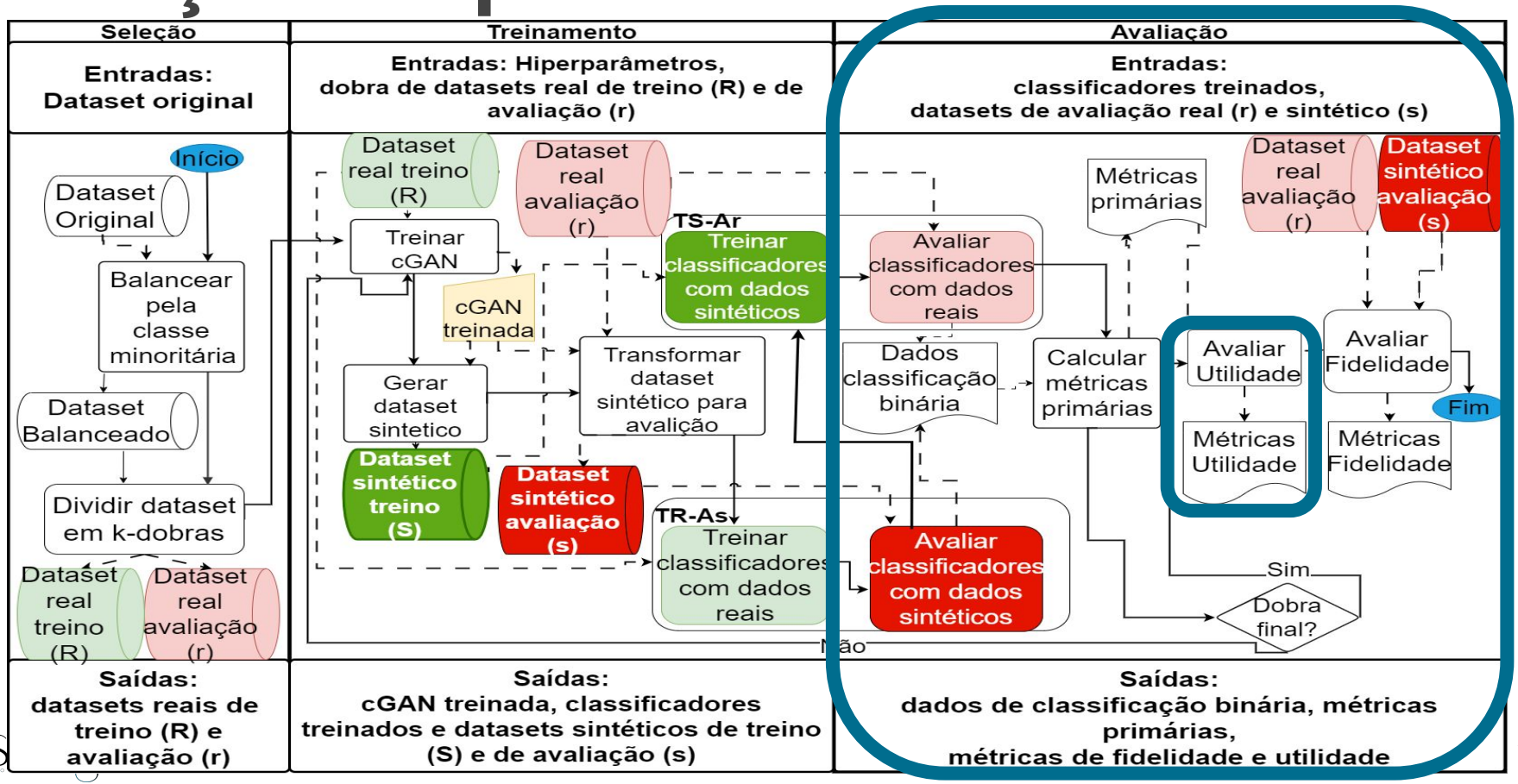


Treinado com real
Avaliado com sintético

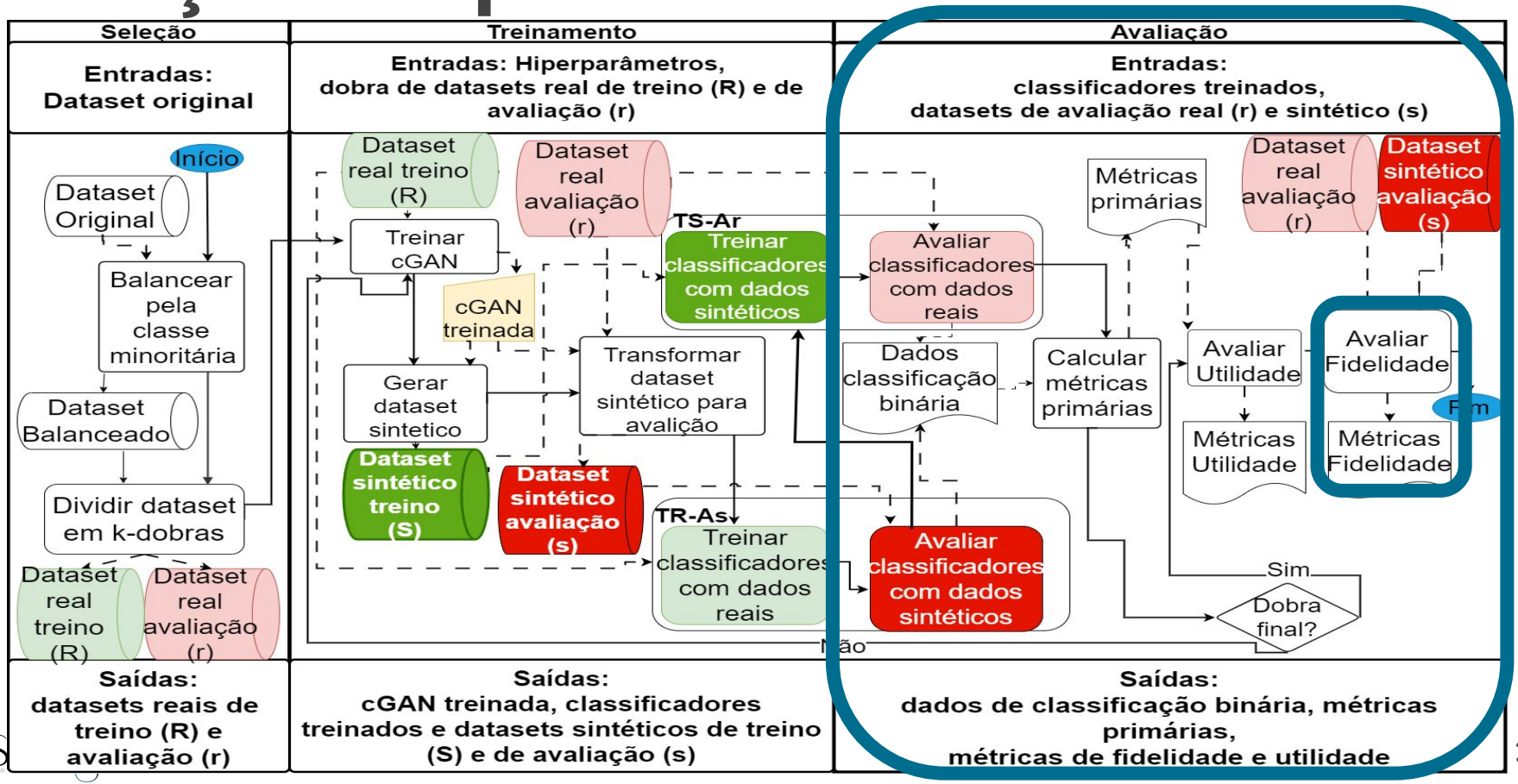
Solução Proposta



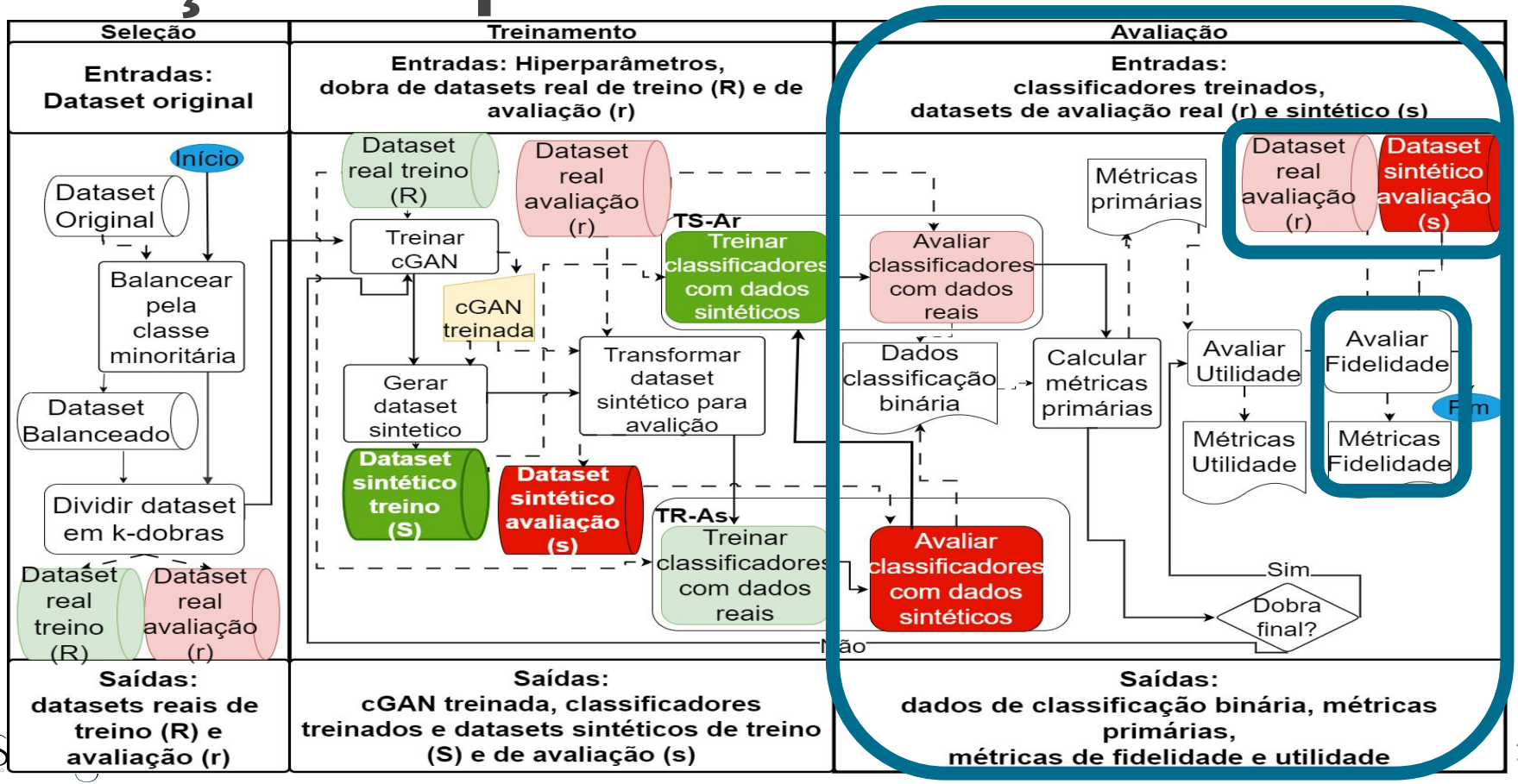
Solução Proposta



Solução Proposta



Solução Proposta



Ferramentas e tecnologias

Ambiente de testes:

- Ubuntu 22.04
- Python:
  - 3.8.10
  - 3.8.2
  - 3.10.12
- Docker 24.0.7 e 20.10.5

Bibliotecas python:

- Numpy 1.21.5
- Keras 2.9.0
- Tensorflow 2.9.1
- Pandas 1.4.4
- Scikit-learn 1.1.1
- Mlflow 2.12.1

Avaliação (Conjuntos de dados)

Dataset	Características	Amostras		
		Malware	Benignos	Total
Kronodroid E	276	50%	50%	20.000
Kronodroid R	285	50%	50%	20.000

Avaliação (Métricas)

- **Fidelidade Estatística:**
  - Erro quadrático médio
  - Valor de p (Teste de Wilcoxon)
  - Similaridade de Cossenos

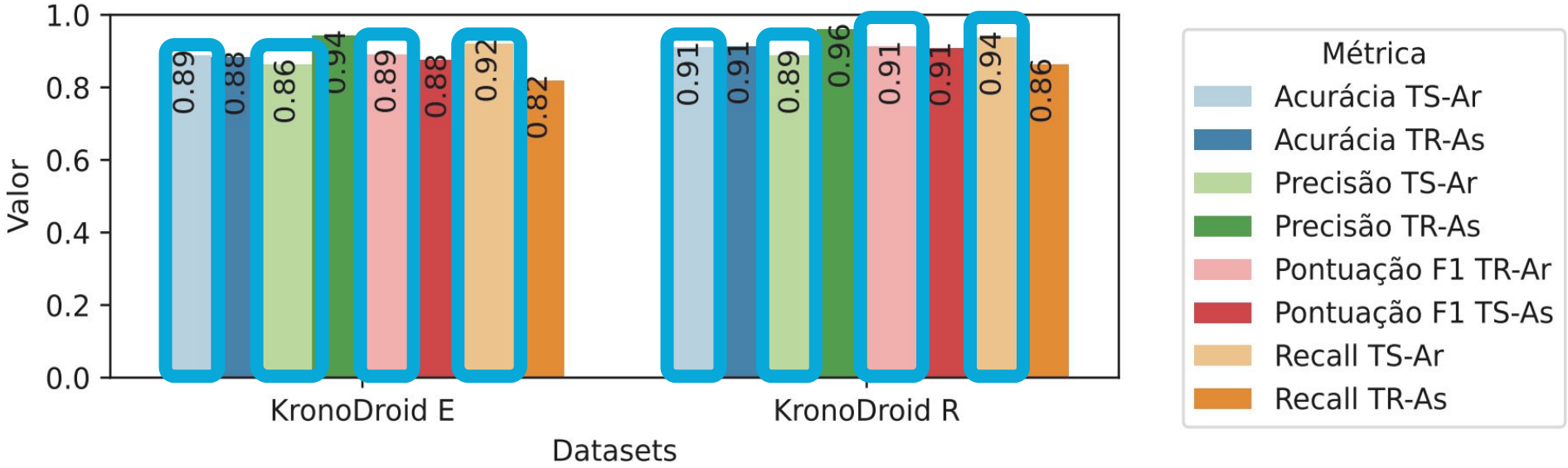
Teste de wilcoxon: Individual comparisons by Ranking Method

Avaliação (Métricas)

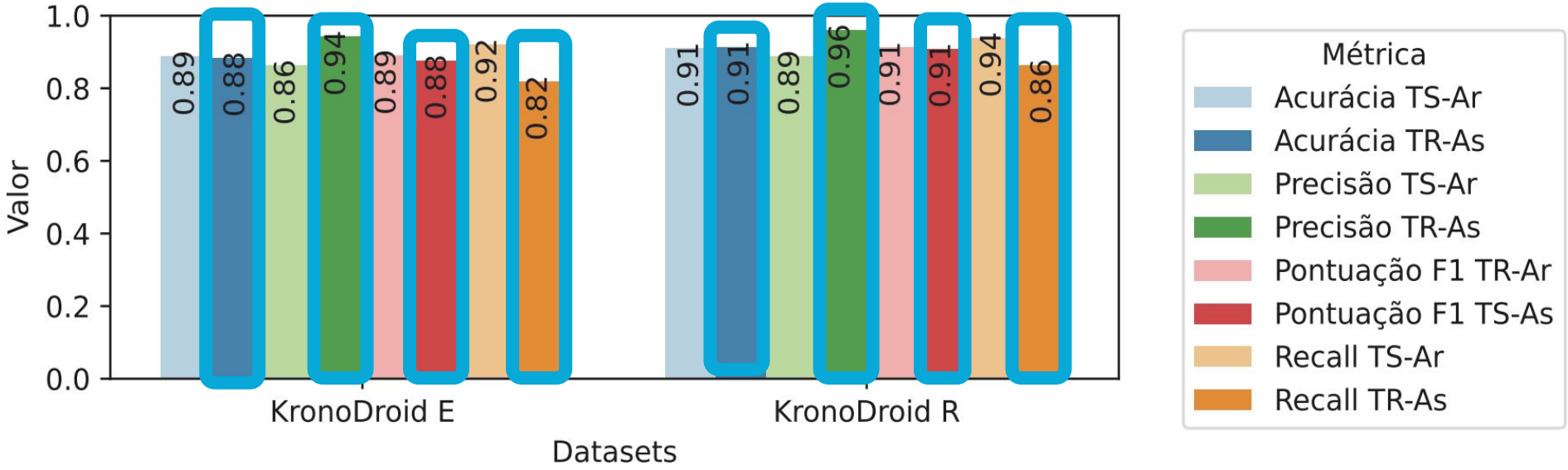
- **Utilidade dos dados:**
  - Acurácia
  - Precisão
  - Recall
  - Pontuação F1

Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs

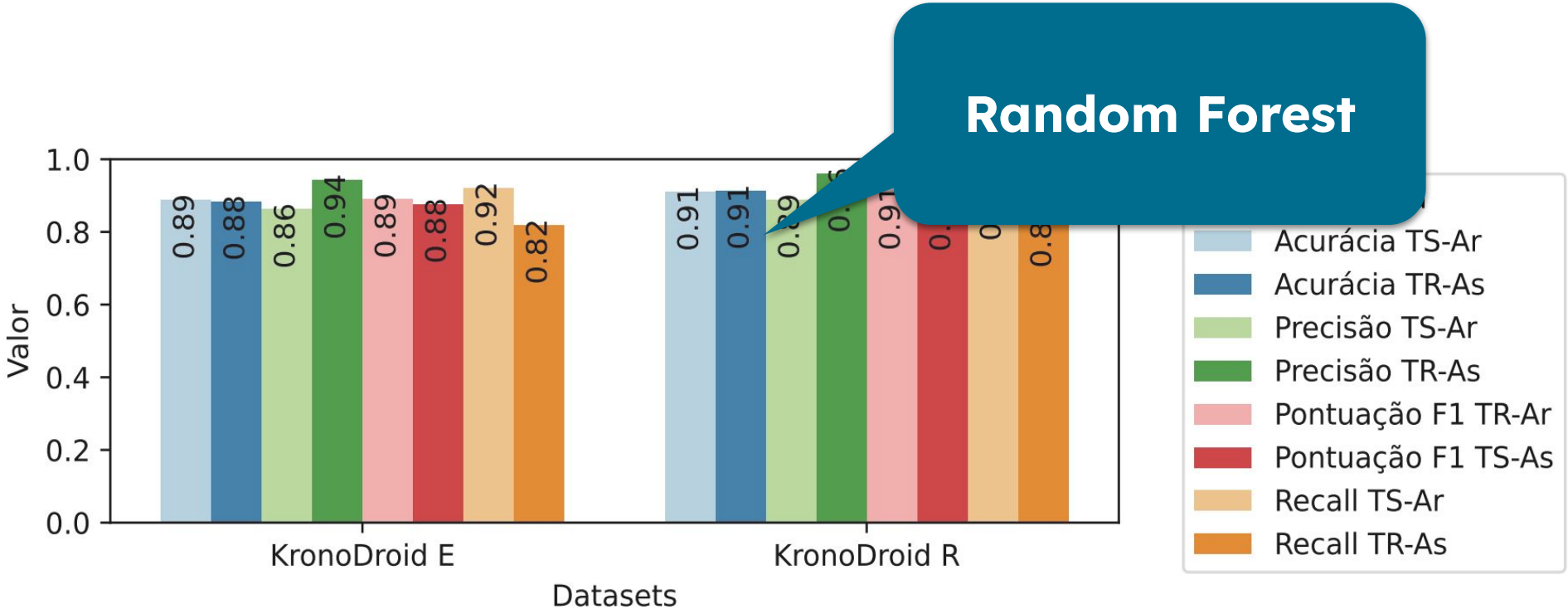
Avaliação (métricas de utilidade)



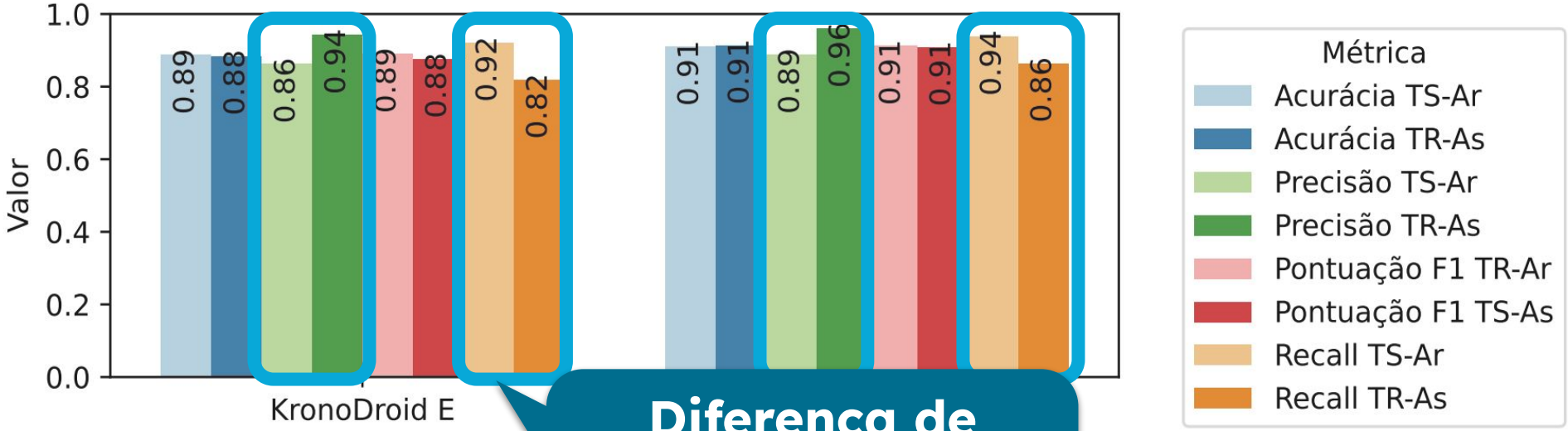
Avaliação (métricas de utilidade)



Avaliação (métricas de utilidade)

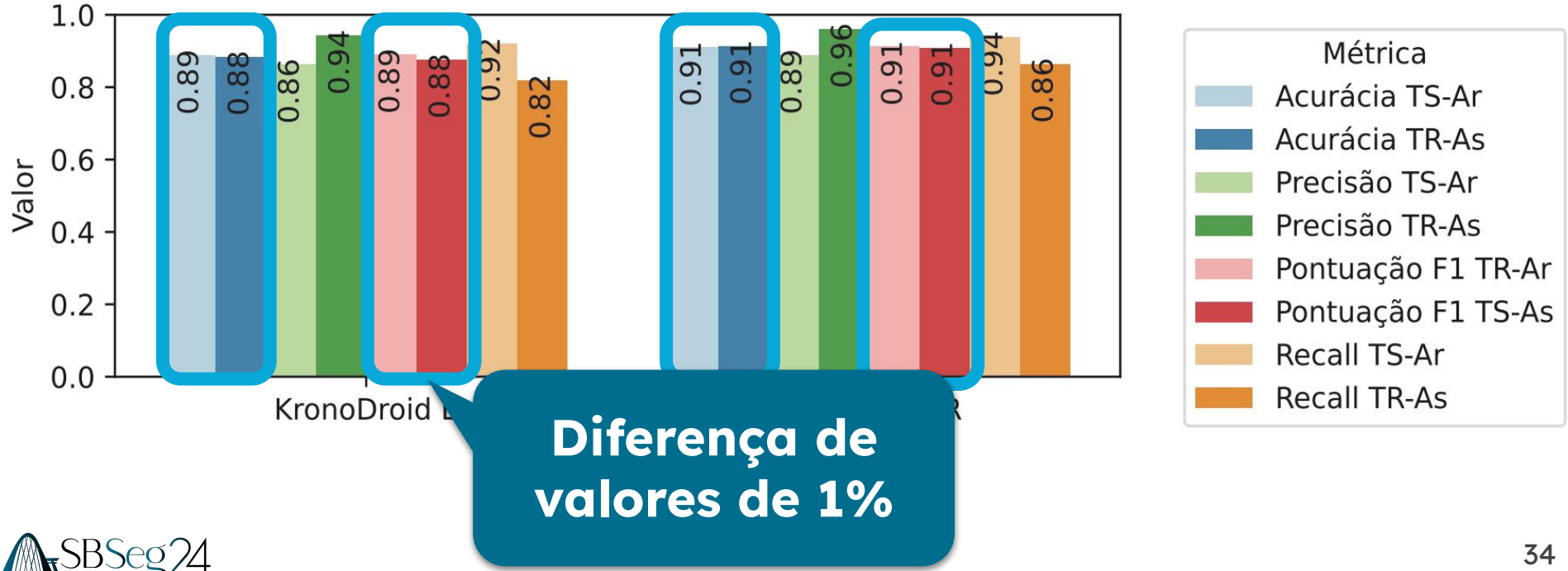


Avaliação (métricas de utilidade)



Diferença de valores entre 8% e 12%

Avaliação (métricas de utilidade)



Avaliação (Métricas de fidelidade)

Dataset	***P value (Random forest)***					
	Acurácia	**Precisão**	**Pontuação F1**	**Recall**	**Média**	**AUC**
Kronodroid E	0,770	0,002	0,064	0,002	0,209	0,375
Kronodroid R	0,922	0,002	0,846	0,002	0,444	0,625

Avaliação (Métricas de fidelidade)

Dataset	*P value (Random forest)*					
	Acurácia	**Precisão**	**Pontuação F1**	**Recall**	**Média**	**AUC**
Kronodroid E	0,770	0,002	0,064	0,002	0,209	0,375
Kronodroid R	0,922	0,002	0,846	0,002	0,444	0,625

Abaixo do limiar de 0,05

Avaliação (Métricas de fidelidade)

Dataset	*P value (Random forest)*					
	Acurácia	**Precisão**	**Pontuação F1**	**Recall**	**Média**	**AUC**
Kronodroid E	0,770	0,002	0,064	0,002	0,209	0,375
Kronodroid R	0,922	0,002	0,846	0,002	0,444	0,625

Acima do limiar de 0,05

Avaliação (Métricas de fidelidade)

Dataset	*P value (Random forest)*					
	Acurácia	**Precis**		**all**	**Média**	**AUC**
Kronodroid E	0,770	0,00			0,209	0,375
Kronodroid R	0,922	0,002	0,846	0,002	0,444	0,625

Média dos valores de *P* do classificador

Avaliação (Métricas de fidelidade)

Dataset	Média *P value*					
	Random forest	**Decision tree**	**Ada boost**	**Perceptron**	**SVM**	**XG boost**
Kronodroid E	0,209	0,060	0,198	0,173	0,234	0,291
Kronodroid R	0,444	0,195	0,125	0,345	0,482	0,173

Avaliação (Métricas de fidelidade)

Todos os classificadores estão acima do limiar de 0,05

dia *P value*

				Perceptron	SVM	XG boost
Kronodroid E	0,209	0,060	0,198	0,173	0,234	0,291
Kronodroid R	0,444	0,195	0,125	0,345	0,482	0,173

Avaliação (Métricas de fidelidade)

Dataset	Positivo		Falso	
	Cosseno	**Erro quadrático**	**Cosseno**	**Erro quadrático**
Krondroid E	0,77	0,10	0,77	0,07
Kronodroid R	0,71	0,11	0,73	0,07

Avaliação (Métricas de fidelidade)

malware

benigno

Dataset	Positivo		Falso	
	Cosseno	**Erro quadrático**	**Cosseno**	**Erro quadrático**
Krondroid E	0,77	0,10	0,77	0,07
Kronodroid R	0,71	0,11	0,73	0,07

Avaliação (Métricas de fidelidade)

Alta similaridade de cosseno

Dataset	Falso			
	Cosseno	Erro quadrático	Cosseno	Erro quadrático
Krondroid E	0,77	0,10	0,77	0,07
Kronodroid R	0,71	0,11	0,73	0,07

Avaliação (Métricas de fidelidade)

Dataset	Positivo		Falso	
	Cosseno	Erro quadrático		Erro quadrático
Krondroid E	0,77	0,10	0,77	0,07
Kronodroid R	0,71	0,11	0,73	0,07

Erro quadrático próximo de 0

Trabalhos relacionados

Trabalho	Métricas	Domínio
Xu and Veeramachaneni, 2018	pontuação F1, acurácia, erro quadrático médio e absoluto	Dados gerais
Amin, M. et. al, 2022	pontuação F1, recall, acurácia e precisão, AUC, FPR e cobertura	Malware Android
Este trabalho	pontuação F1, recall, acurácia, precisão, AUC, cosseno de similaridade, erro quadrático, valor de p e discrepância média	Malware Android

Considerações Finais

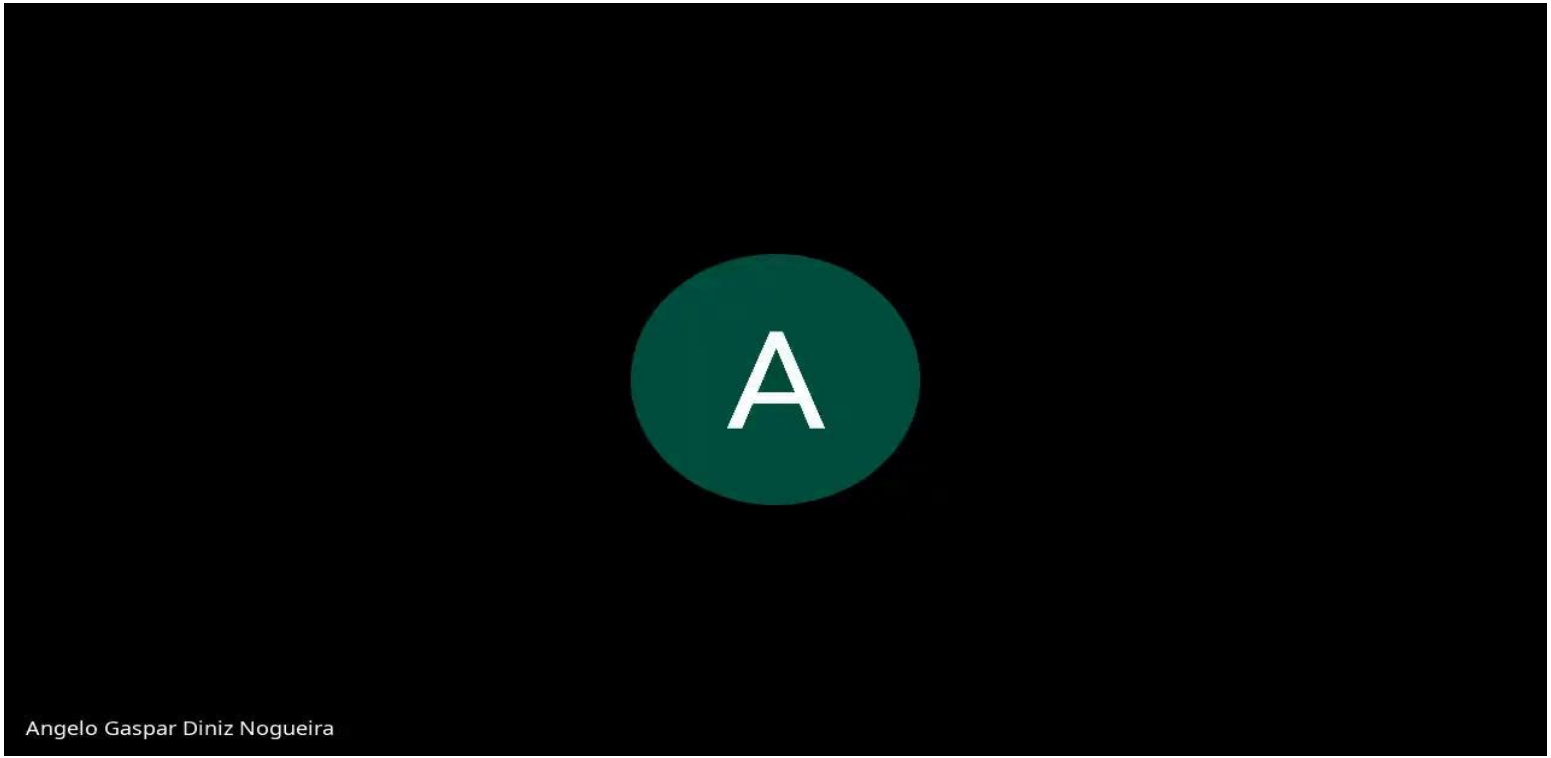
- MalSynGen disponível no GitHub:

- MalSynGen produz dados:
  - Fiéis em relação aos dados originais
  - Úteis para diversos classificadores

Link para o GitHub

Considerações Finais

Demonstração



Trabalhos futuros

- Incluir outras métricas
- Ampliar o número de datasets
- Uma análise comparativa de desempenho com ferramentas similares (gerais)
- Generalizar a ferramenta para incluir outros métodos de geração de dados e outras potenciais aplicações

Obrigado!

Landing page: https://malwaredatalab.github.io/