# Avaliação de Métodos de Seleção de Características de Amostras Android com a Ferramenta FS3E (v2)

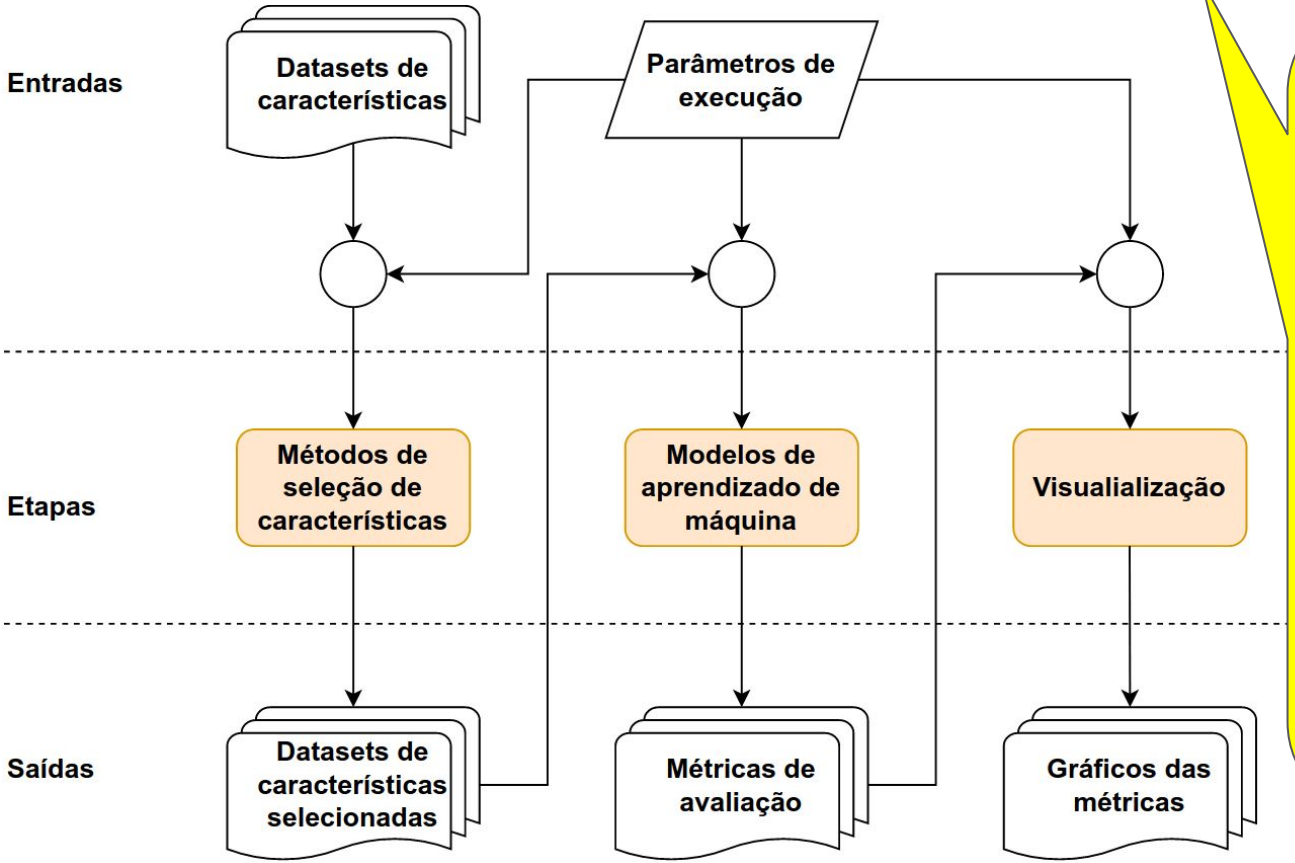

UFAM

**Nicolas Neves**, Vanderson Rocha, Diego Kreutz, Hendrio Bragança, Eduardo Feitosa

unipampa

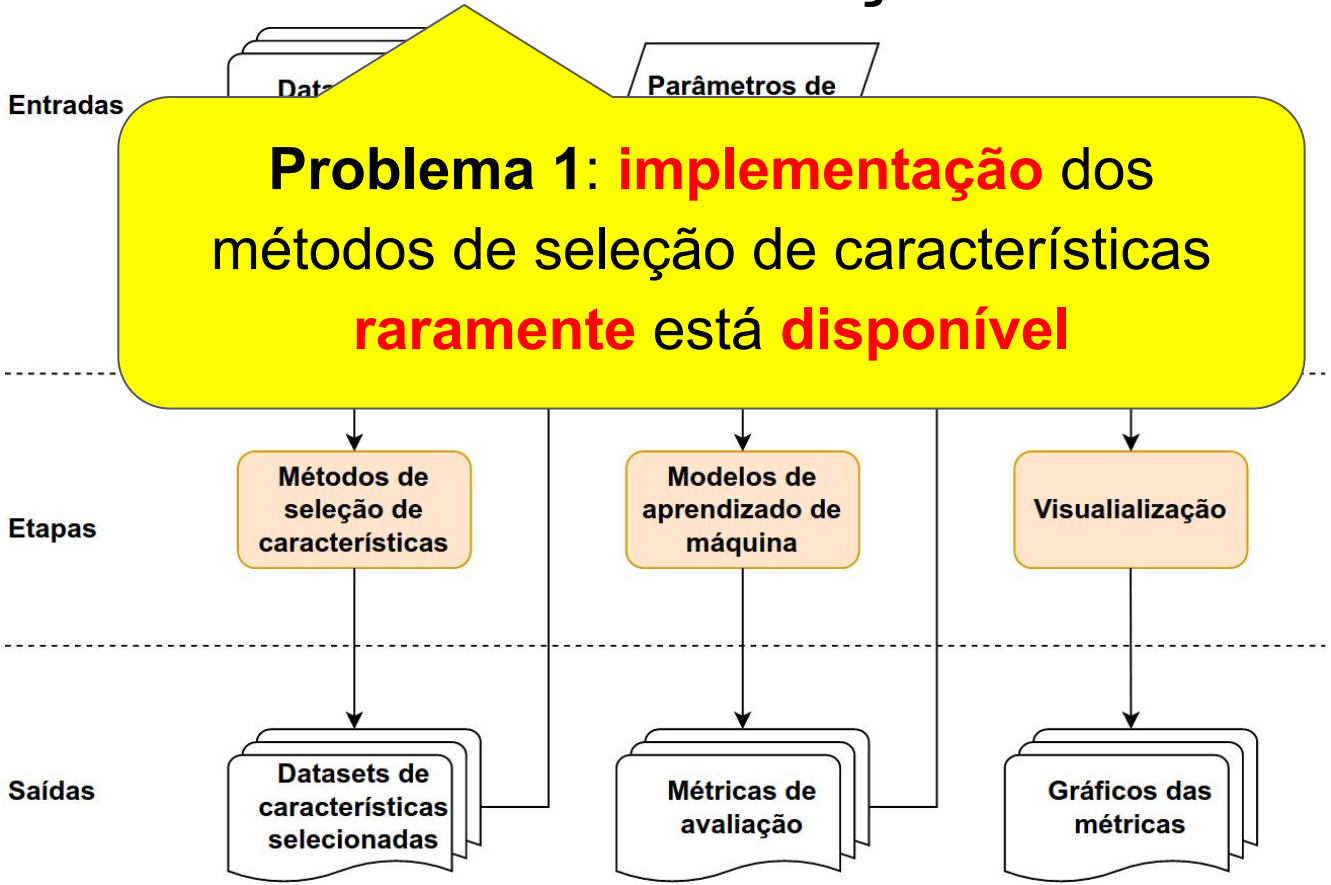
Universidade Federal do Pampa

# O Framework FS3E

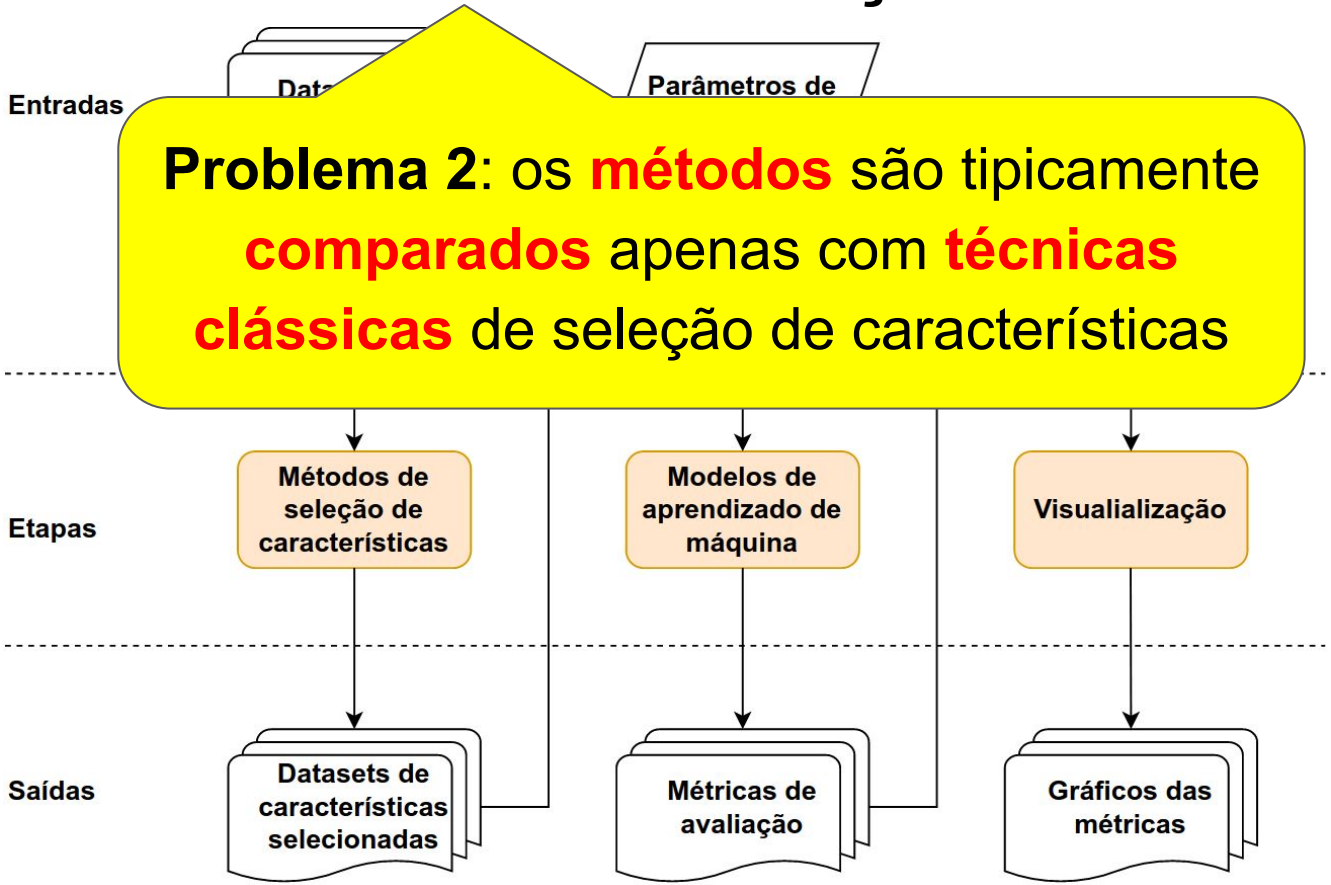



Um framework para **implementação**, **disponibilização** e **sistematização** da **avaliação** de métodos de seleção de características.

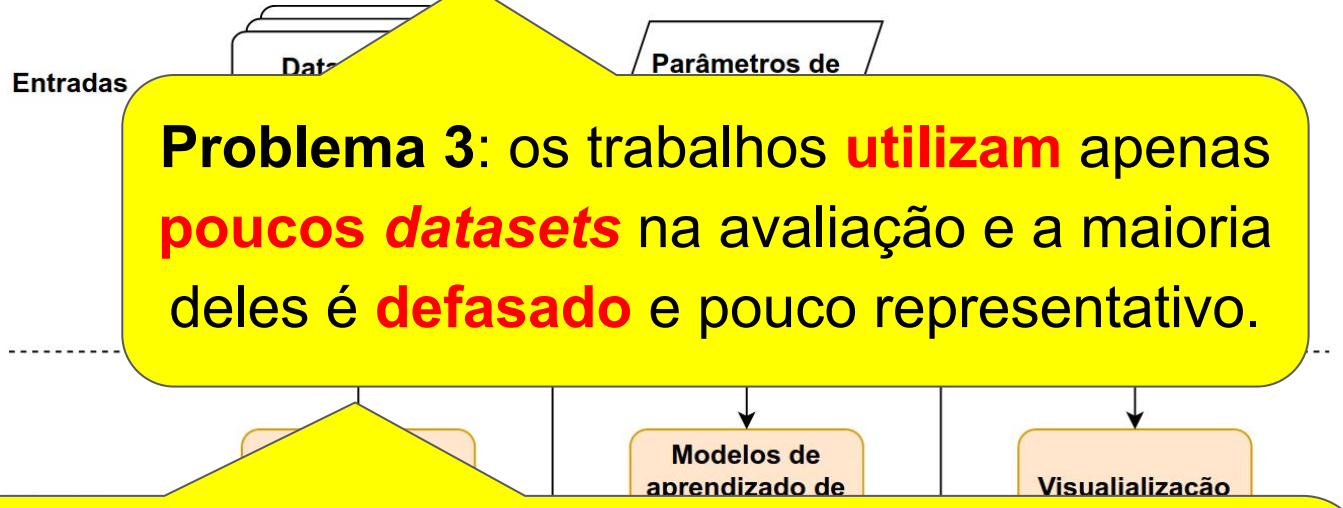
# FS3E: motivação



# FS3E: motivação



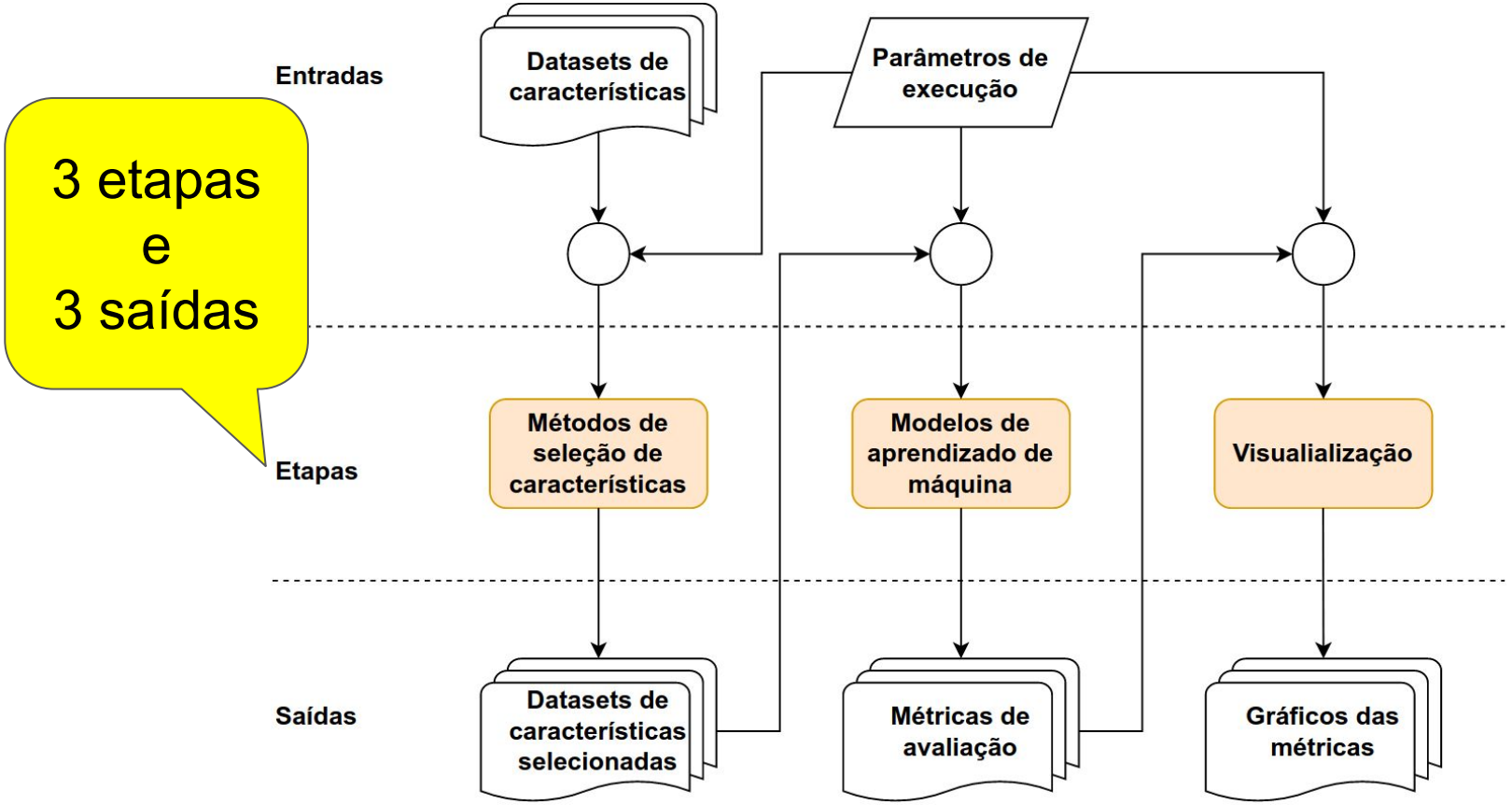
# FS3E: motivação




**Problema 3**: os trabalhos **utilizam** apenas **poucos *datasets*** na avaliação e a maioria deles é **defasado** e pouco representativo.

Exemplo (**MT**): dataset (D) com 11449 amostras (5279 B e 6170 M).
**Malwares**: VirusTotal, VirusShare e Drebin.
Apps **benignas**: Google Play Store

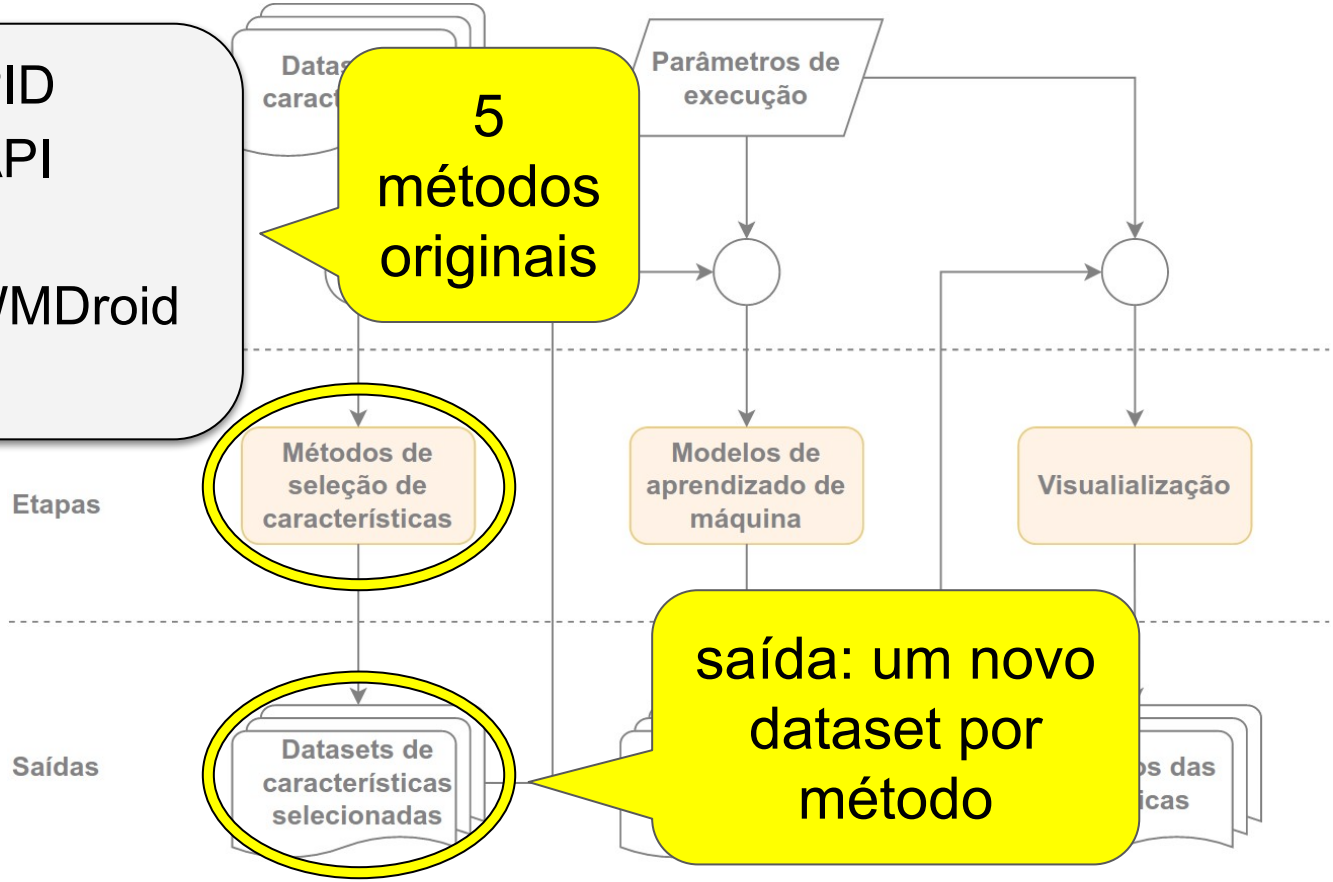
# FS3E: etapas e fluxo de dados



# FS3E: etapa 1

- SigPID
- SigAPI
- RFG
- JOWMDroid
- LR

5 métodos originais



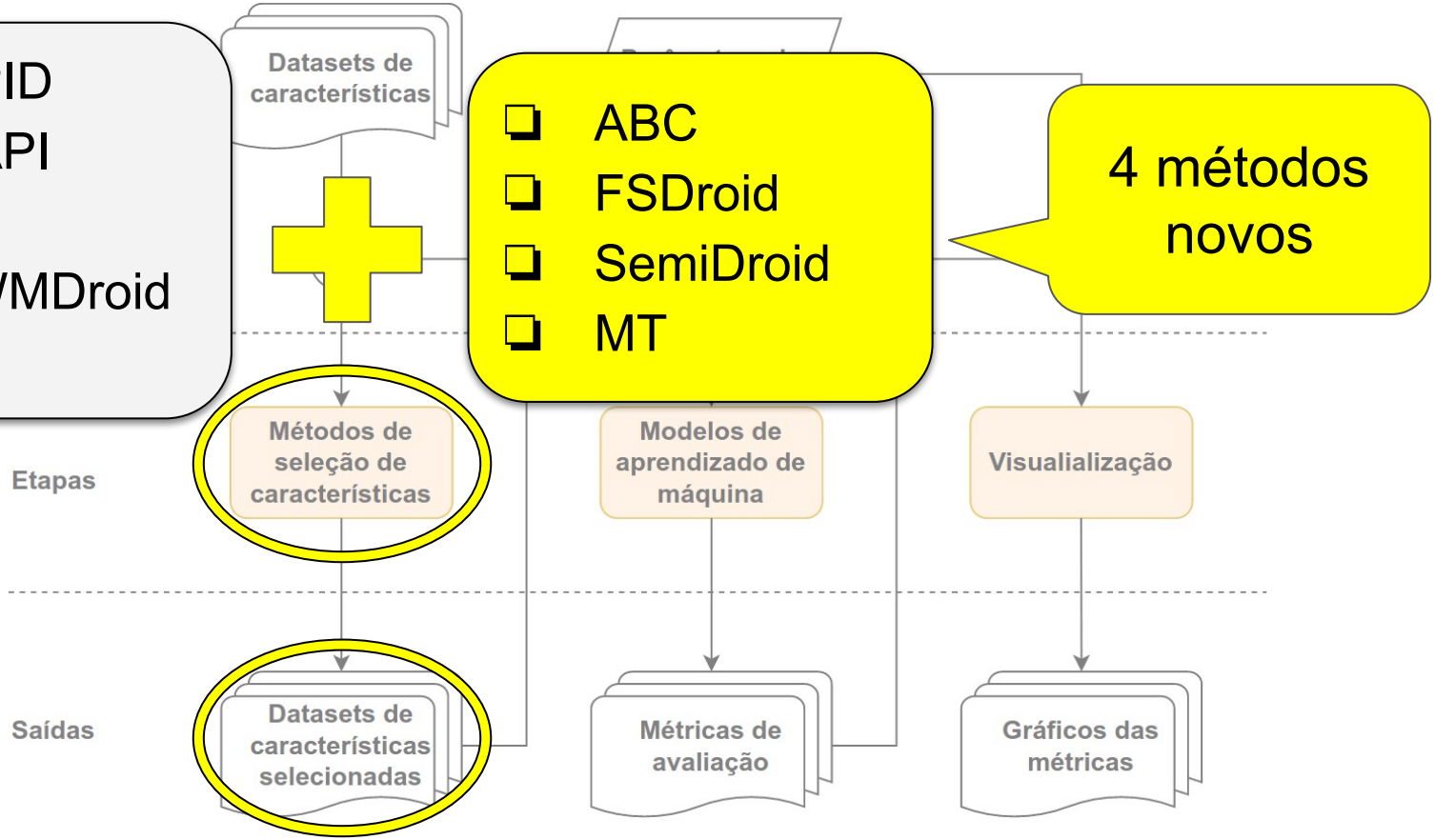

saída: um novo dataset por método

# FS3E: etapa 1

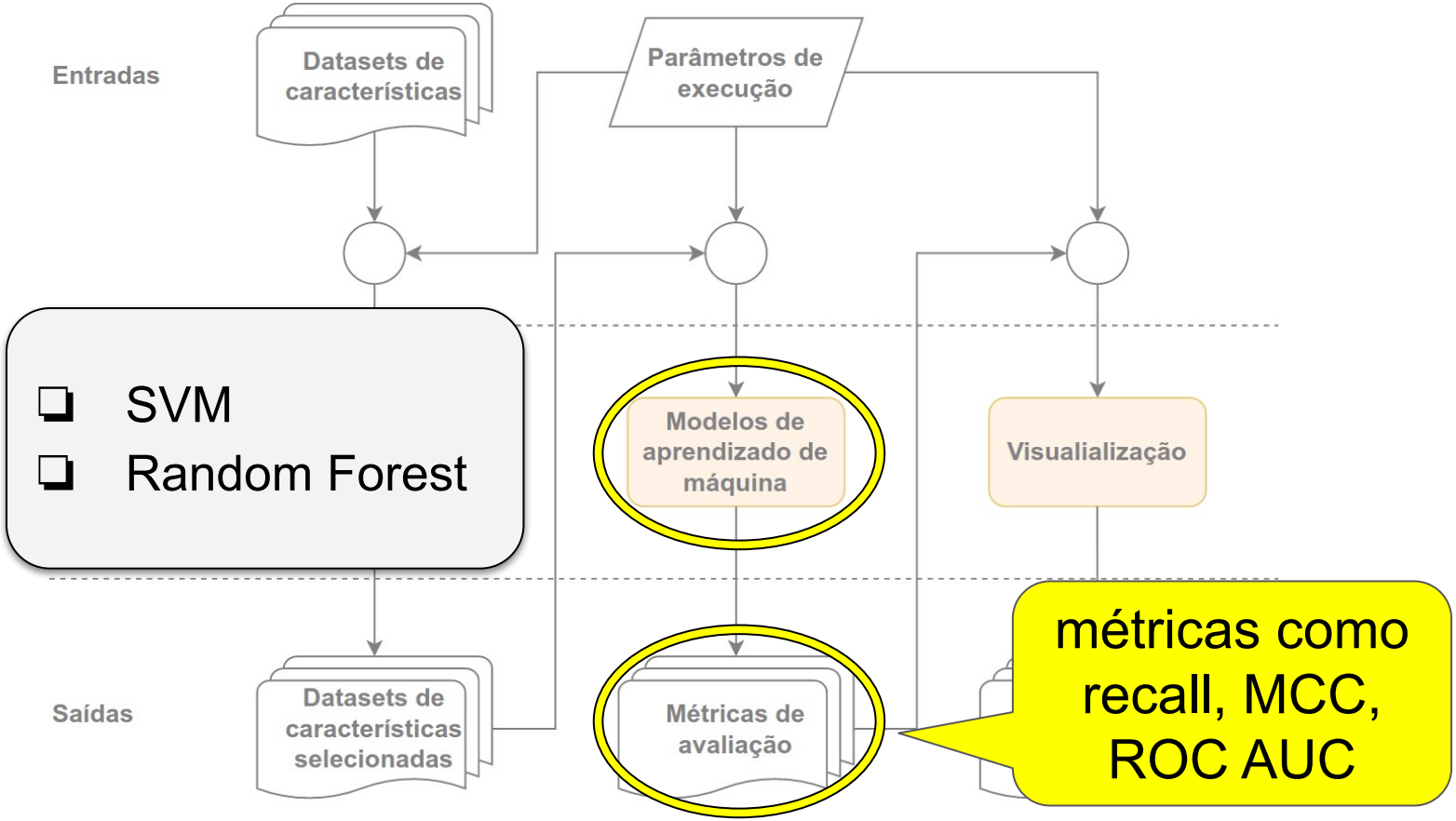
- SigPID
- SigAPI
- RFG
- JOWMDroid
- LR

- ABC
- FSDroid
- SemiDroid
- MT

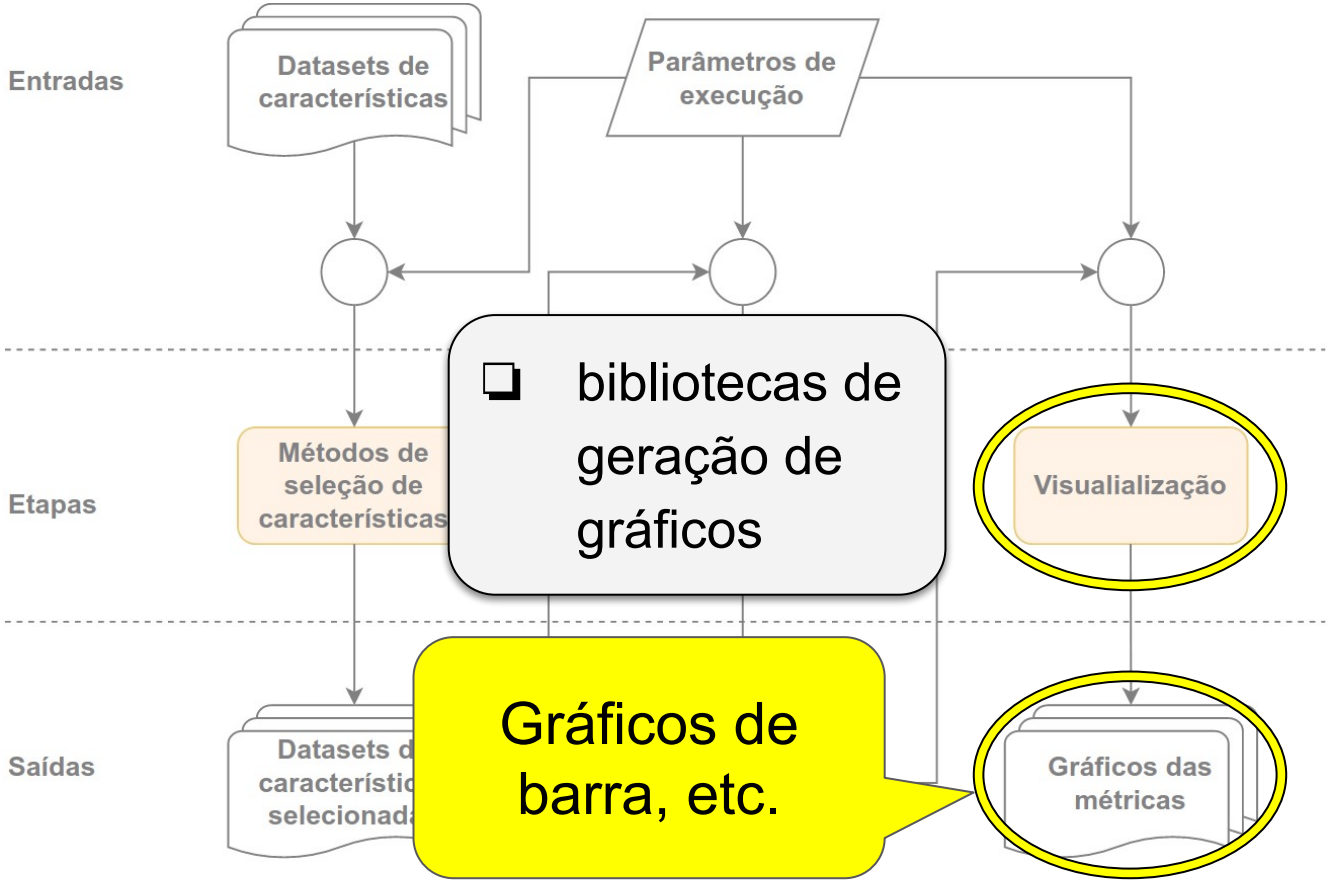
4 métodos novos



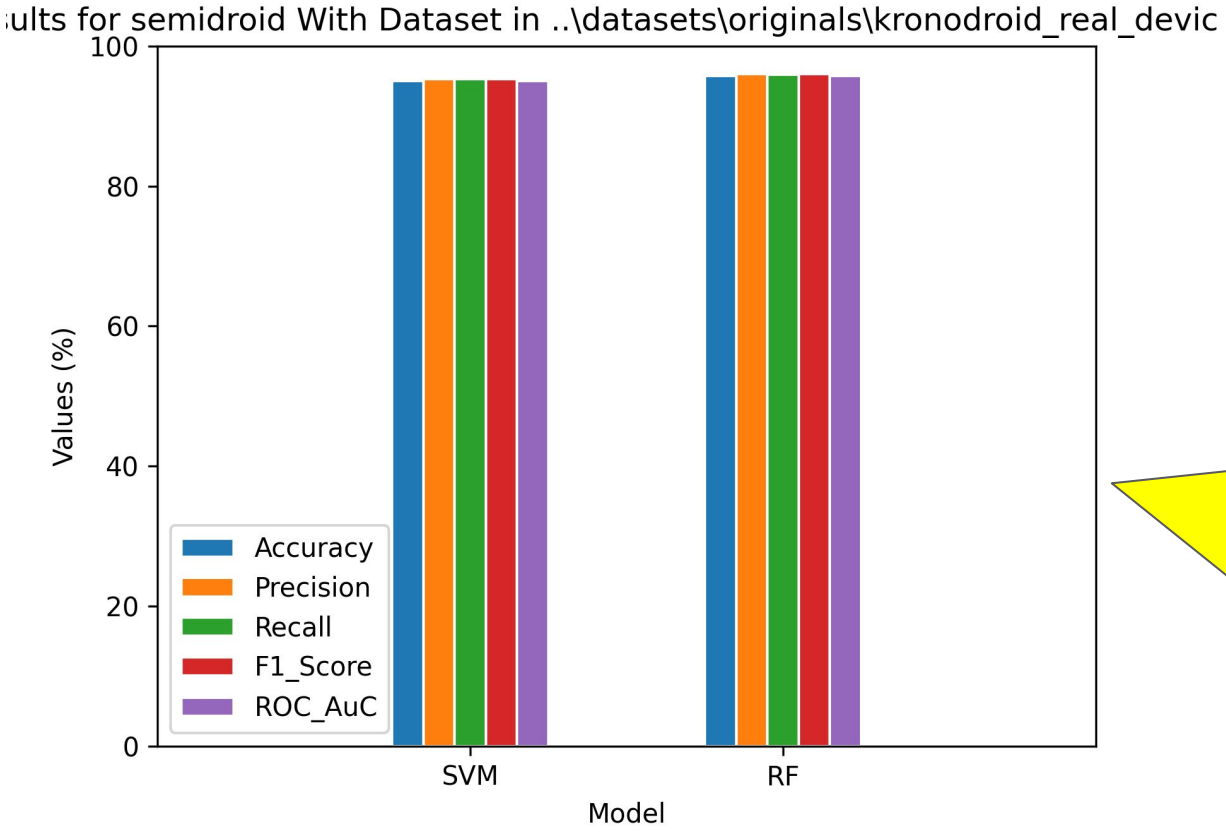
# FS3E: etapa 2


Entradas
Datasets de características
Parâmetros de execução
SVM
Random Forest
Modelos de aprendizado de máquina
Visualialização
Saídas
Datasets de características selecionadas
Métricas de avaliação
métricas como recall, MCC, ROC AUC

# FS3E: etapa 3


Entradas
Datasets de características
Parâmetros de execução
Etapas
Métodos de seleção de características
Visualialização
Saídas
Datasets d
característic
selecionad
Gráficos das métricas
❑ bibliotecas de geração de gráficos
Gráficos de barra, etc.

# FS3E: etapa 3




Exemplo de gráfico de métricas do SVM e RF para o dataset KronoDroid

# Dados dos *Datasets* Originais

| Dataset | Tipo Carac. | Amostras | Cars. | % M / B |
|---|---|---|---|---|
| Adroit | P | 11476 | 166 | 29.8 / 70.2 |
| Androcrawl | P, I, CAPI | 96732 | 141 | 10.5 / 89.5 |
| Android_Permissions | P | 26864 | 151 | 66.21 / 33.79 |
| DefeseDroid | P, I | 11975 | 2938 | 50.1 / 49.9 |
| Drebin-215 | P, CAPI, CS, I | 15036 | 215 | 37 / 63 |
| KronoDroid | P, CS | 78137 | 383 | 53 / 47 |

# Dados dos *Datasets* Originais

| Dataset | Tipo Carac. | Amostras | Cars. | % M / B |
|---|---|---|---|---|
| Adroit | **P** | 11476 | 166 | 29.8 / 70.2 |
| Androcrawl | **P, I, CAPI** | | | |
| Android_Permissions | **P** | | | |
| DefeseDroid | **P, I** | | | |
| Drebin-215 | **P, CAPI, CS, I** | | | |
| KronoDroid | **P, CS** | 78137 | 383 | 53 / 47 |

Tipos de características:

**P** = **P**ermissões

**I** = **I**ntenções

**CAPI** = **C**hamadas de **API**

**CS** = **C**hamadas de **S**istema

# Dados dos *Datasets* Originais

| Dataset | Tipo Carac. | Amostras | Cars. | % M / B |
|---|---|---|---|---|
| **Adroit** | P | **11476** | 166 | 29.8 / 70.2 |
| **Androcrawl** | P, I, CAPI | **96732** | **141** | 10.5 / 89.5 |
| Android_Permissions | P | 26864 | 151 | 66.21 / 33.79 |
| **DefeseDroid** | P, I | 11975 | **2938** | 50.1 / 49.9 |
| | | 15036 | 215 | 37 / 63 |
| | | 78137 | 383 | 53 / 47 |

Variação significativa no número de amostras e características.

# Dados dos *Datasets* Originais

| Dataset | Tipo Carac. | Amostras | Cars. | % M / B |
|---|---|---|---|---|
| Adroit | P | 11476 | 166 | 29.8 / 70.2 |
| **Androcrawl** | | | 11 | **10.5 / 89.5** |
| **Android_Permissions** | | | | **66.21 / 33.79** |
| **DefeseDroid** | | | 8 | **50.1 / 49.9** |
| Drebin-215 | | | 5 | 37 / 63 |
| KronoDroid | P, CS | 78137 | 383 | 53 / 47 |

*Datasets* **desbalanceados**
*Datasets* com desbalanceamento **inverso**
*Datasets* **balanceados**

# *Dataset* Adroit

| Dataset | Recall | MCC | C AUC | % Redu. |
|---|---|---|---|---|
| Ori | | | | 0.0 |
| A | | | | 0.60 |
| FSI | | | | 82.63 |
| JOW | | | | 71.86 |
| I | | | | 39.52 |
| R | | | | 26.95 |
| SemiDroid | 69.89 | 0.72 | 83.28 | 89.22 |
| MT | 71.01 | 0.69 | 82.58 | 97.01 |
| SigPID | 67.64 | 0.74 | 82.98 | 97.01 |

# amostras: **11476**

# características: **166**

Tipo(s): **Permissões**

Proporção: **29.8% M / 70.2% B**

# *Dataset* Adroit

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 79.72 | 0.59 | 81.27 | 0.0 |
| ABC | 79.75 | 0.60 | 81.30 | |
| FSDroid | 70.22 | 0.73 | 83.69 | |
| JOWNDroid | 1.96 | 0.04 | 50.46 | |
| LR | 73.00 | 0.72 | 84.25 | |
| RFG | 76.95 | 0.67 | 83.5 | |
| SemiDroid | 69.89 | 0.72 | 83.28 | |
| **MT** | **71.01** | **0.69** | **82.58** | **97.01** |
| **SigPID** | **67.64** | **0.74** | **82.98** | **97.01** |

MT e SigPID tem excelente taxa de redução e métricas boas

# *Dataset* Adroit

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 79.72 | 0.59 | 81.27 | 0.0 |
| ABC | 79.75 | 0.60 | | |
| FSDroid | 70.22 | 0.73 | | |
| JOWNDroid | 1.96 | 0.04 | | |
| LR | 73.00 | 0.72 | | |
| **RFG** | **76.95** | **0.67** | **83.55** | **26.95** |
| SemiDroid | 69.89 | 0.72 | 83.28 | 89.22 |
| MT | 71.01 | 0.69 | 82.58 | 97.01 |
| SigPID | 67.64 | 0.74 | 82.98 | 97.01 |

RFG consegue um recall muito bom a uma taxa de redução próxima de 30%

# *Dataset* Androcrawl

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Origi | | | | 0.0 |
| A | | | | 1.41 |
| FSI | | | | 79.58 |
| JOW | | | | 78.17 |
| I | | | | 94.37 |
| R | | | | 84.51 |
| Sem | | | | 88.73 |
| MT | | | | 96.48 |
| SigPID | 87.01 | 0.89 | 93.19 | 95.77 |

# amostras: **96732**

# características: **141**

Tipo(s): **Permissões, *Intents*, Chamadas de API**

Proporção: **10.5% M / 89.5% B**

# *Dataset* Androcrawl

FSDroid, LR, SemiDroid e SigPID com excelente taxa de redução e métricas boas

| | | | | edu. |
|---|---|---|---|---|
| | | | | 0 |
| ABC | .16 | 0.91 | 94.28 | 1.41 |
| **FSDroid** | **89.76** | **0.91** | **94.56** | **79.58** |
| JOWNDroid | 1.03 | 0.03 | 50.34 | 78.17 |
| **LR** | **87.10** | **0.89** | **93.19** | **94.37** |
| RFG | 25.68 | 0.42 | 62.44 | 84.51 |
| **SemiDroid** | **86.51** | **0.90** | **93.01** | **88.73** |
| MT | 4.72 | 0.11 | 51.97 | 96.48 |
| **SigPID** | **87.01** | **0.89** | **93.19** | **95.77** |

# *Dataset* Androcrawl

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | | | | 0.0 |
| ABC | | | | 1.41 |
| FSDroid | | | | 79.58 |
| JOWNDroid | | | | 78.17 |
| LR | | | | 94.37 |
| RFG | 25. | .42 | 62.44 | 84.51 |
| SemiDroid | 86.51 | 0.90 | 93.01 | 88.73 |
| **MT** | **4.72** | **0.11** | **51.97** | **96.48** |
| SigPID | 87.01 | 0.89 | 93.19 | 95.77 |

O *threshold* de 0.8 do MT acaba **selecionando características menos relevantes** para *datasets* muito **desbalanceados** (1M para 9B)

# *Dataset* Android_Permissions

| Dataset | Recall | MCC | AUC | % Redu. |
|---|---|---|---|---|
| Ori | | | | 0.0 |
| A | | | | .32 |
| FS | | | | 1.05 |
| JOW | | | | 4.61 |
| | | | | 4.61 |
| R | | | | 5.53 |
| Ser | | | | 3.82 |
| MI | | | | 97.37 |
| SigPID | 98.93 | 0.08 | 51.11 | 71.71 |

# amostras: **26864**

# características: **151**

Tipo(s): **Permissões**

**Proporção**: **66.21% M / 33.79% B**

# *Dataset* Android_Permissions

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 89.76 | 0.15 | **55.67** | 0.0 |
| | | .15 | **55.24** | 1.32 |
| | | .13 | **53.63** | 71.05 |
| J | | .06 | **51.37** | 54.61 |
| | | | **51.45** | 54.61 |
| | | .04 | **50.61** | 85.53 |
| SemiDroid | 95.48 | 0.10 | **52.56** | 88.82 |
| MT | 100.00 | 0.00 | **50.00** | 97.37 |
| SigPID | 98.93 | 0.08 | **51.11** | 71.71 |

*Dataset* com **desbalanceamento inverso** e **características compostas**

# *Dataset* DefenseDroid

| Dataset | Recall | MCC | AUC | % Redu. |
|---|---|---|---|---|
| Ori | | | | 0.0 |
| A | | | | .09 |
| FS | | | | 3.89 |
| JOW | | | | 4.97 |
| | | | | 9.08 |
| R | | | | 9.24 |
| Ser | | | | 9.96 |
| MT | | | | 99.79 |
| SigPID | 86.12 | 0.80 | 89.84 | 97.57 |

# amostras: **11975**

# características: **2938**

Tipo: **Permissões, *Intents***

**Proporção: 50.1% M / 49.9% B**

# *Dataset* DefenseDroid

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 91.00 | 0.85 | 92.39 | 0.0 |
| ABC | 90.70 | 0.84 | 91.98 | 3.09 |
| **FSDroid** | 85.05 | 0.77 | 88.57 | **98.89** |
| JOWNDroid | 57.50 | 0.50 | 73.57 | [illegible].97 |
| LR | | | | |
| RFG | | | | |
| SemiDroid | 89.23 | 0.84 | 92.11 | [illegible].96 |
| **MT** | 86.27 | 0.63 | 81.40 | **99.79** |
| **SigPID** | 86.12 | 0.80 | 89.84 | **97.57** |

*Dataset* **balanceado** leva a **taxas agressivas** de **redução**

# *Dataset* DefenseDroid

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 91.00 | 0.85 | 92.39 | 0.0 |
| ABC | 90.70 | 0.84 | 91.98 | 3.09 |
| **FSDroid** | **85.05** | **0.77** | **88.57** | **98.89** |
| JOWNDroid | 57.50 | 0.50 | 73.57 | 54.97 |
| | | | | 29.08 |
| | | | | 99.24 |
| | | | | 89.96 |
| **MT** | 86.27 | 0.63 | **81.40** | 99.79 |
| **SigPID** | **86.12** | **0.80** | **89.84** | **97.57** |

FSDroid e SigPID conseguem taxa de redução acima de 97% e taxas de classificação muito boas

# *Dataset* Drebin-215

| Dataset | Recall | MCC | AUC | % Redu. |
|---|---|---|---|---|
| Ori | | | | 0.0 |
| A | | | | 0.0 |
| FS | | | | 0.74 |
| JOW | | | | 4.26 |
| | | | | 5.19 |
| R | | | | 0.56 |
| Ser | | | | 9.35 |
| | | | | 0.28 |
| SigF | | | | 90.28 |

# amostras: **15036**

# características: **215**

Tipo(s): **Permissões, *Intents*, Chamadas de API, Chamadas de Sistema**

Proporção: **37% M / 63% B**

# *Dataset* Drebin-215

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 97.50 | 0.97 | 98.53 | 0.0 |
| ABC | 97.50 | 0.97 | 98.53 | 0.0 |
| FSDroid | 88.03 | 0.87 | 92.59 | 90.74 |
| **JOWNDroid** | **59.01** | **0.56** | **75.73** | **84.26** |
| LR | | 0.86 | 92.45 | 85.19 |
| RF | | | | 80.56 |
| SemiD | | | | 89.35 |
| MT | | | | 90.28 |
| SigPID | 90.55 | 0.89 | 93.98 | 90.28 |

JOWNDroid obteve um péssimo desempenho

# *Dataset* KronoDroid

| Dataset | Recall | MCC | AUC | % Redu. |
|---|---|---|---|---|
| Ori | | | | 0.0 |
| A | | | | .39 |
| FS | | | | 3.85 |
| JOW | | | | 7.35 |
| | | | | 4.11 |
| R | | | | 2.33 |
| Ser | | | | 9.55 |
| MT | | | | 96.52 |
| SigPID | 93.69 | 0.87 | 93.59 | 93.73 |

# amostras: **78137**

# características: **383**

Tipo: **Permissões, Chamadas de Sistema**

**Proporção: 53% M / 47% B**

# *Dataset* KronoDroid

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 97.30 | 0.95 | 97.62 | 0.0 |
| ABC | 97.17 | 0.95 | 97.48 | 1.39 |
| FSDroid | 93.92 | 0.85 | 92.57 | 88.85 |
| **JOWNDroid** | **93.07** | **0.13** | **54.11** | **77.35** |
| LR | 94.20 | 0.87 | 63 | 64.11 |
| RF | | | | 92.33 |
| SemiD | | | | 89.55 |
| MT | | | | 96.52 |
| SigPID | 93.69 | 0.87 | 93.59 | 93.73 |

Novamente, JOWNDroid obteve um péssimo desempenho

# *Dataset* KronoDroid

| Dataset | Recall | MCC | ROC AUC | % Redu. |
|---|---|---|---|---|
| Original | 97.30 | 0.95 | 97.62 | 0.0 |
| ABC | 97.17 | 0.95 | 97.48 | 1.39 |
| FSDroid | 93.92 | 0.85 | 92.57 | 88.85 |
| JOWNDroid | 93.07 | 0.13 | 54.11 | 77.35 |
| LR | | | | 64.11 |
| RF | | | | 92.33 |
| SemiD | | | | 89.55 |
| MT | 89.49 | 0.76 | .93 | 96.52 |
| **SigPID** | **93.69** | **0.87** | **93.59** | **93.73** |

Novamente, SigPID consegue **excelentes resultados** para ***datasets* balanceados**

# Métodos mais estáveis (FSDroid)

| Dataset | Recall | MCC | ROC AUC | # Car. | % R. |
|---|---|---|---|---|---|
| Adroit | | 3 | **83.69** | 29 | **82.63** |
| Androcraw | | | **94.56** | 29 | **79.58** |
| DefenseDro | | 7 | **88.57** | 32 | **98.89** |
| Drebin-21 | | 7 | **92.59** | 20 | **90.74** |
| KronoDroid | | 35 | **92.57** | 32 | **88.85** |

Altas taxas de **redução** e taxas de **classificação** muito boas

# Métodos mais estáveis (SemiDroid)

| Dataset | Recall | MCC | ROC AUC | # Car. | % R. |
|---|---|---|---|---|---|
| Adroit | | | **83.28** | 18 | **89.22** |
| Androcrawl | | | **93.01** | 16 | **88.73** |
| Defensedroid | | | **92.11** | 289 | **89.96** |
| Drebin-215 | | | **93.91** | 23 | **89.35** |
| KronoDroid | | | **95.66** | 30 | **89.55** |

Altas taxas de **redução** e taxas de **classificação** muito boas

# Considerações Finais

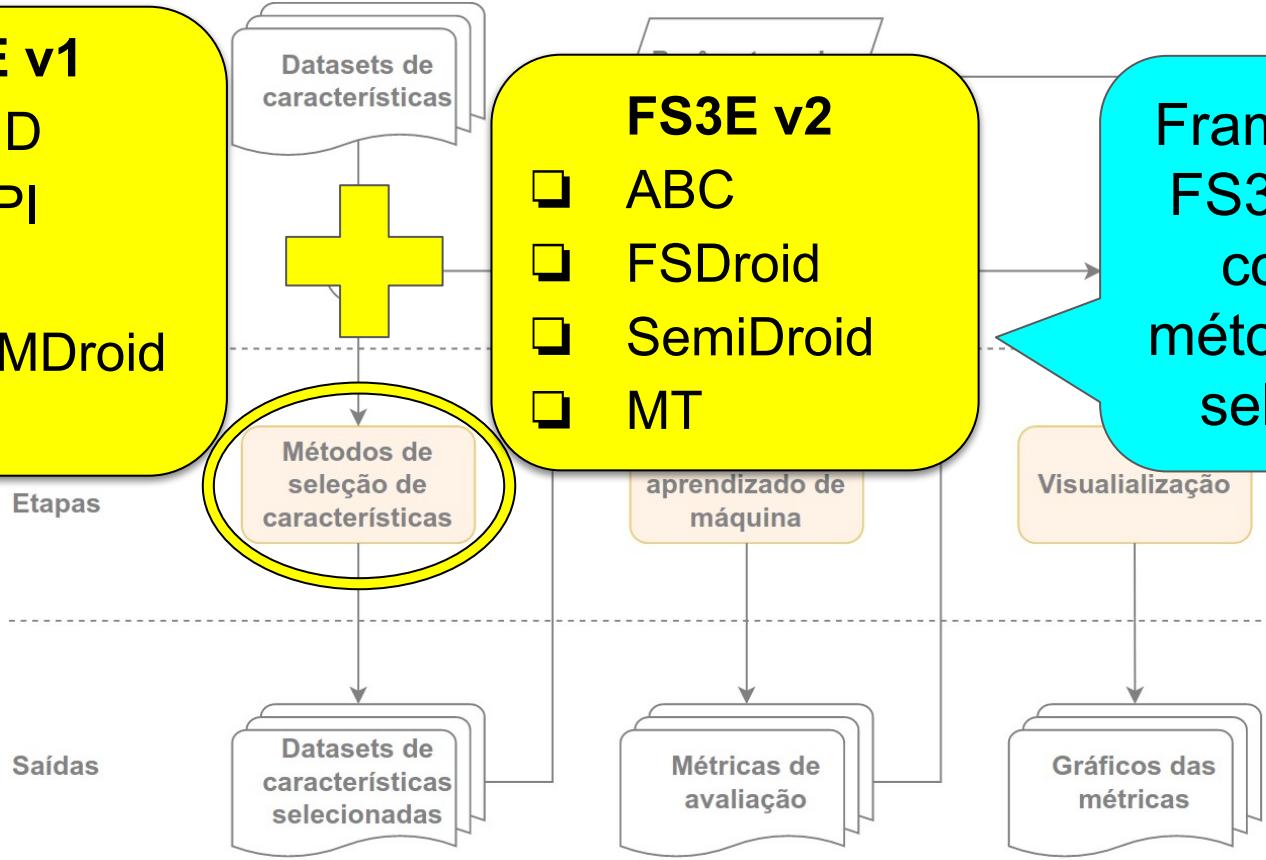
**FS3E v1**

- ❑ SigPID
- ❑ SigAPI
- ❑ RFG
- ❑ JOWMDroid
- ❑ LR

**FS3E v2**

- ❑ ABC
- ❑ FSDroid
- ❑ SemiDroid
- ❑ MT

Framework FS3E (v2) com 9 métodos de seleção



# Considerações Finais

- *Insights* importantes sobre métodos de seleção
- FSDroiod e SemiDroid exibem **maior estabilidade** e capacidade de **generalização**
- RFG e MT **performam mal** em *datasets* com **características** elaboradas por **especialistas**

# Trabalhos Futuros

- Incorporação de mais métodos de seleção de características
- Inclusão de datasets modernos e atualizados
- Utilização de técnicas de explicabilidade (XAI)
- Avaliar os métodos em ambientes de tempo real
- Investigar efeito de ataques adversariais

# Obrigado!

Avaliação de Métodos
com FS3E (v2)

GitHub
FS3E (v2)