Classificação de Malwares

Ciência de Dados para Segurança

Glenda Proença Train

Valfride Wallace do Nascimento

Introdução

- **Smartphones** são parte essencial da vida de grande parte da população
  - Acesso a internet
  - Comunicação entre pessoas
  - Movimentações bancárias
  - ...

- **Malwares**
  - Abusar de vulnerabilidades de segurança
  - Adquirir informações pessoais de seus usuários sem consentimento

Conjunto de Dados

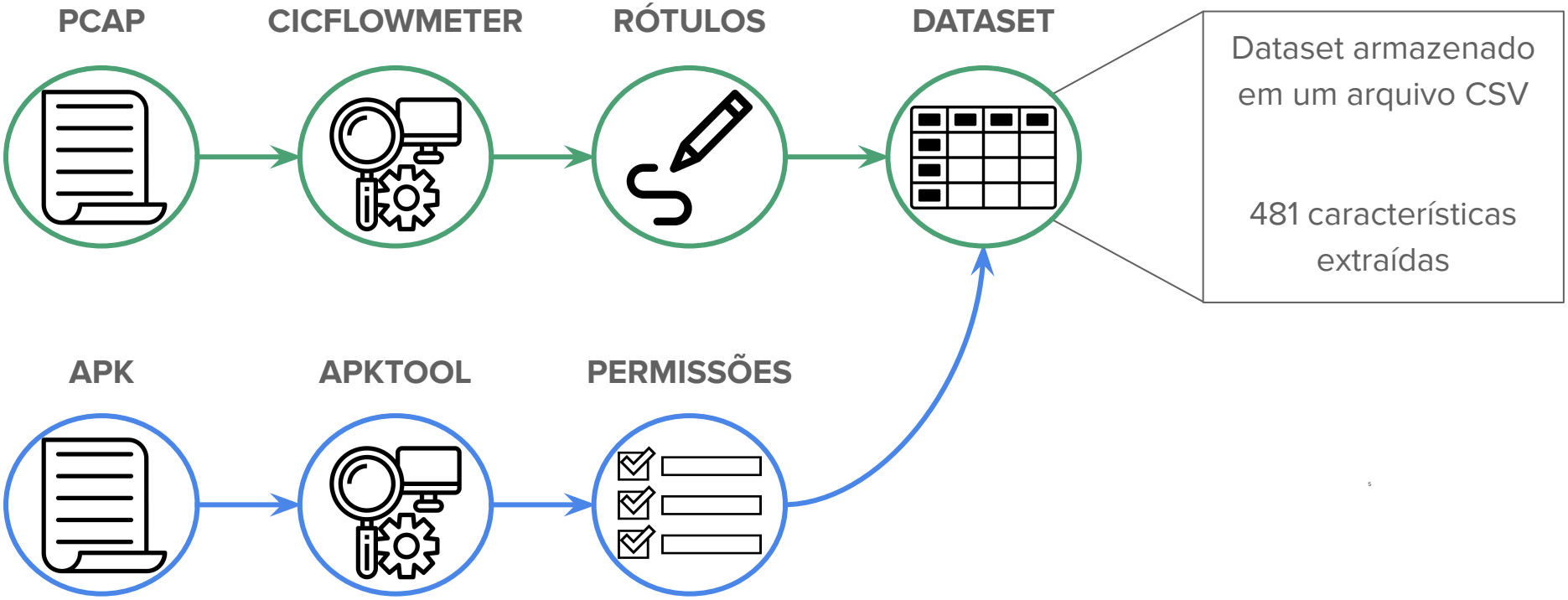
- **Dataset Escolhido:**
  - CIC-AndMal2017
  - Universidade de New Brunswick (UNB)
  - Informações sobre Malwares coletadas de smartphones reais com Android

- **Classes:**
  - Adware
  - Ransomware
  - Scareware
  - SMS Malware
  - Benigna

Objetivo

- Aplicar o processo de **Ciência de Dados**:

Extração de Dados



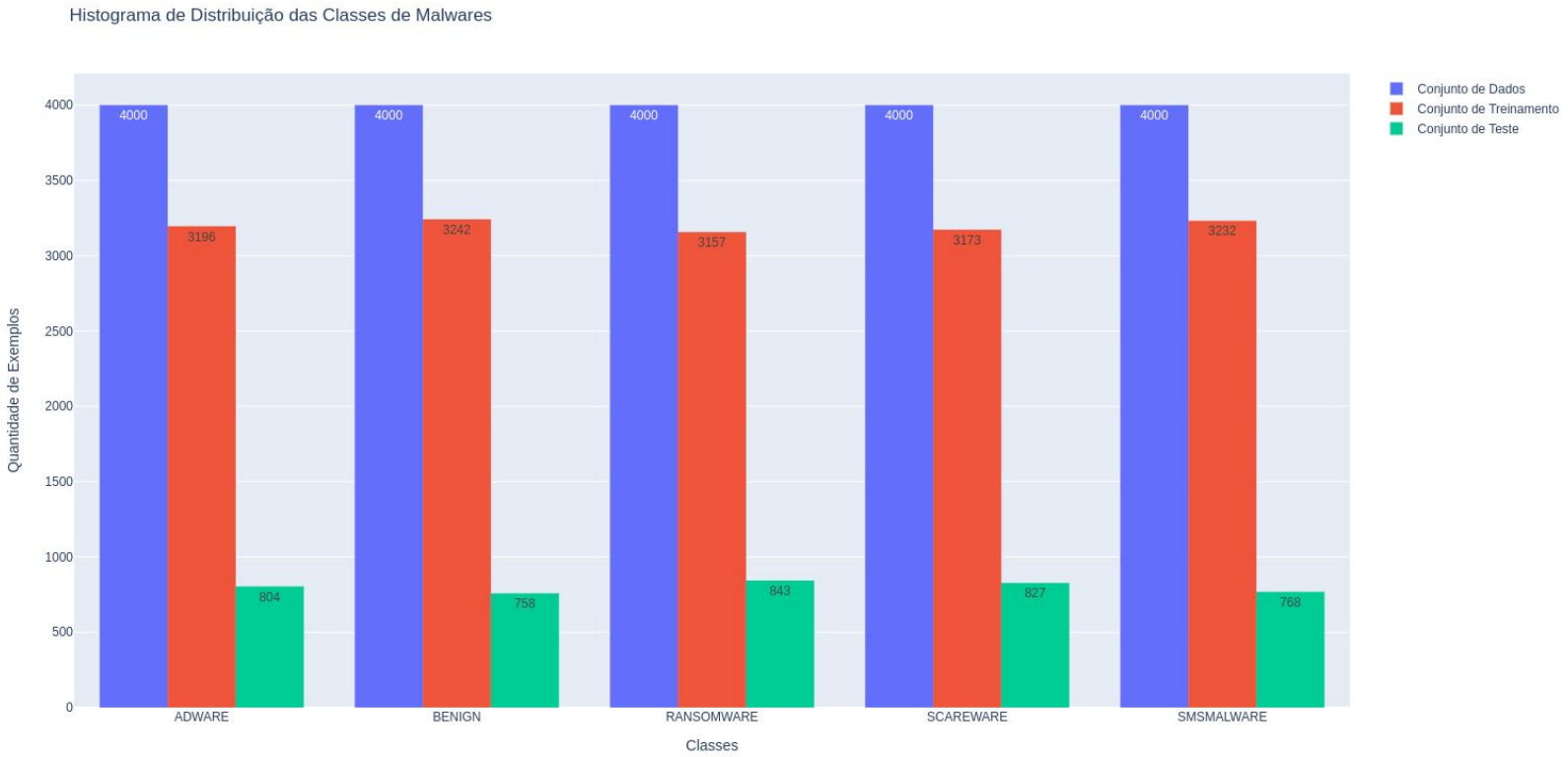
Distribuição das Classes

- **Problema:**
  - Desbalanceamento entre classes
  - 212.084 amostras da classe SMS Malware
  - 902.583 amostras da classe Benigna

- **Solução:**
  - 4.000 amostras aleatórias de cada classe selecionadas
  - *Dataset* final com 20.000 amostras

- **Divisão dos dados:**
  - Conjunto de Treinamento | 80% dos dados
  - Conjunto de Testes | 20% dos dados

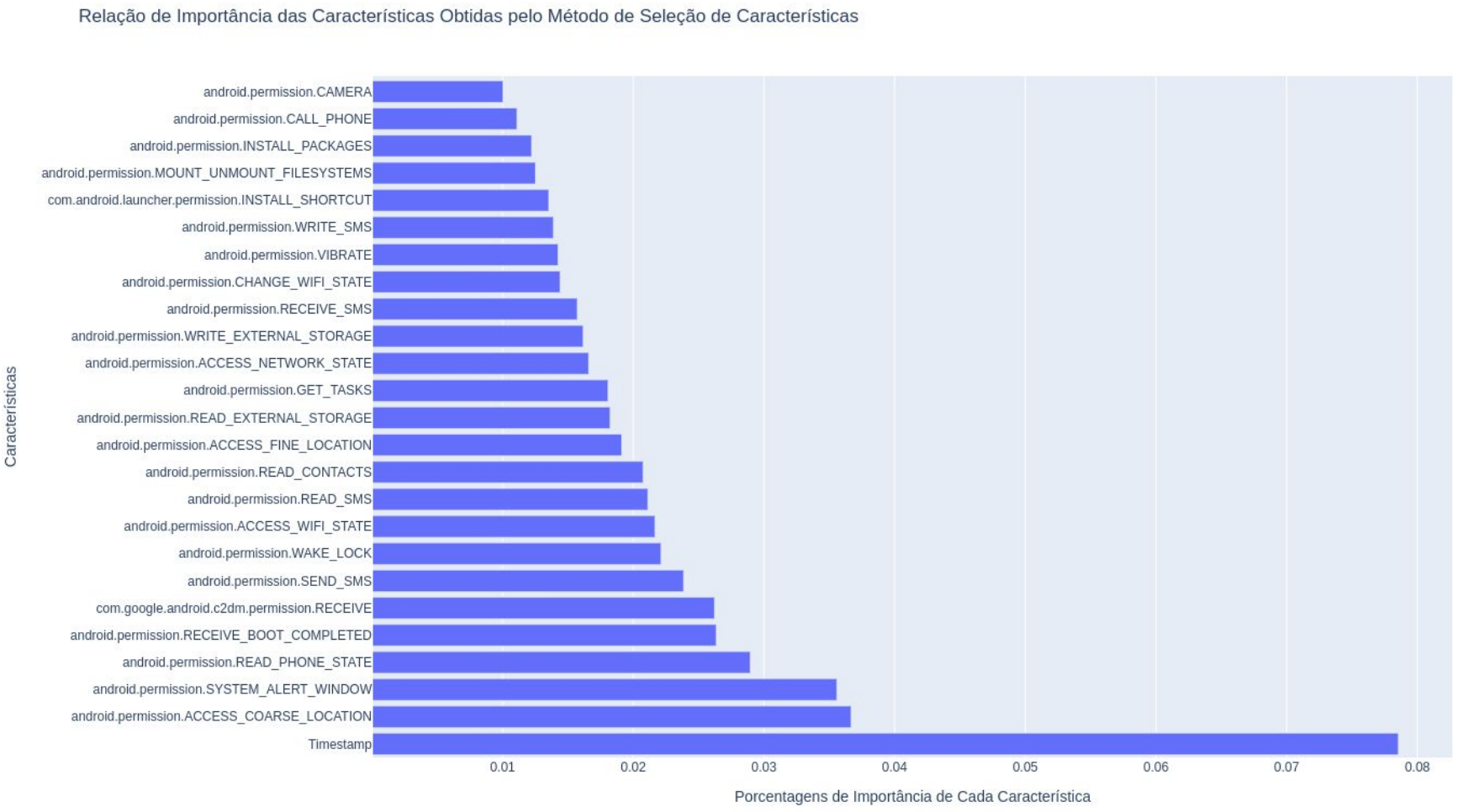
Distribuição das Classes



Modelos Baseados em Aprendizado de Máquina

Etapas de Pré-Processamento

- Transformação das características em **valores numéricos**:
  - **Data** ➔ Timestamp (Pandas | Python)
  - **IP** ➔ Inteiro (Socket Lib | Python)
  - Escala **MinMax** nas características não-booleanas
- **Seleção de Características:**
  - Correlações entre características
  - 481 ➔ 50
  - Extra Trees (*Extremely Randomized Trees Classifier*)
  - Conjunto de validação (4.000 amostras do Conjunto de Treinamento)
- **Resultado:**
  - Somente **2** características do **CICFlowMeter** permaneceram (Data e IP de Origem)
  - As **48** restantes são compostas por **permissões**

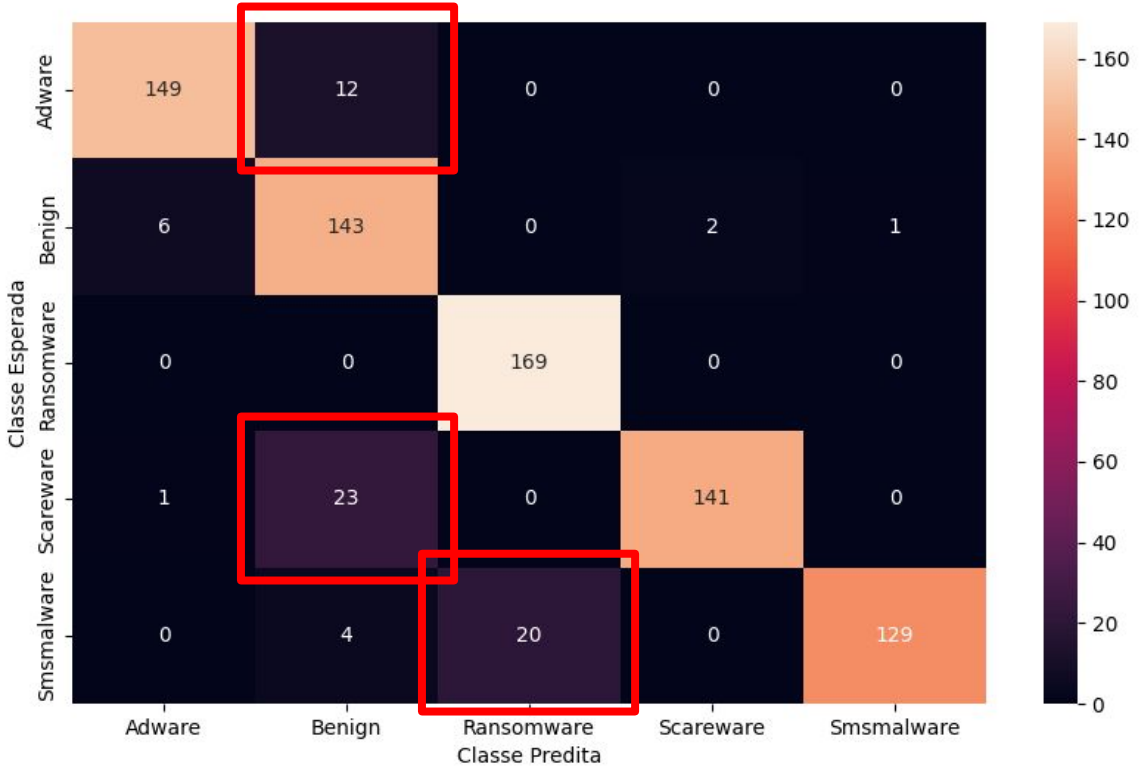

Relação de Importância das Características Obtidas pelo Método de Seleção de Características
android.permission.CAMERA
android.permission.CALL_PHONE
android.permission.INSTALL_PACKAGES
android.permission.MOUNT_UNMOUNT_FILESYSTEMS
com.android.launcher.permission.INSTALL_SHORTCUT
android.permission.WRITE_SMS
android.permission.VIBRATE
android.permission.CHANGE_WIFI_STATE
android.permission.RECEIVE_SMS
android.permission.WRITE_EXTERNAL_STORAGE
android.permission.ACCESS_NETWORK_STATE
android.permission.GET_TASKS
android.permission.READ_EXTERNAL_STORAGE
android.permission.ACCESS_FINE_LOCATION
android.permission.READ_CONTACTS
android.permission.READ_SMS
android.permission.ACCESS_WIFI_STATE
android.permission.WAKE_LOCK
android.permission.SEND_SMS
com.google.android.c2dm.permission.RECEIVE
android.permission.RECEIVE_BOOT_COMPLETED
android.permission.READ_PHONE_STATE
android.permission.SYSTEM_ALERT_WINDOW
android.permission.ACCESS_COARSE_LOCATION
Timestamp
Características
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Porcentagens de Importância de Cada Característica

Classificadores

- **Otimização dos Hiperparâmetros:**
  - Grid Search com Validação Cruzada ➜ **5 pastas**
  - 4.000 amostras do Conjunto de Treinamento
- **Avaliação:**
  - **Validação Cruzada** nos dados do Conjunto de Teste ➜ **5 pastas**
- **Métricas:**
  - Acurácia
  - Precisão
  - Recall
  - F1-Score
  - AUC
  - Curvas ROC
  - Matrizes de Confusão

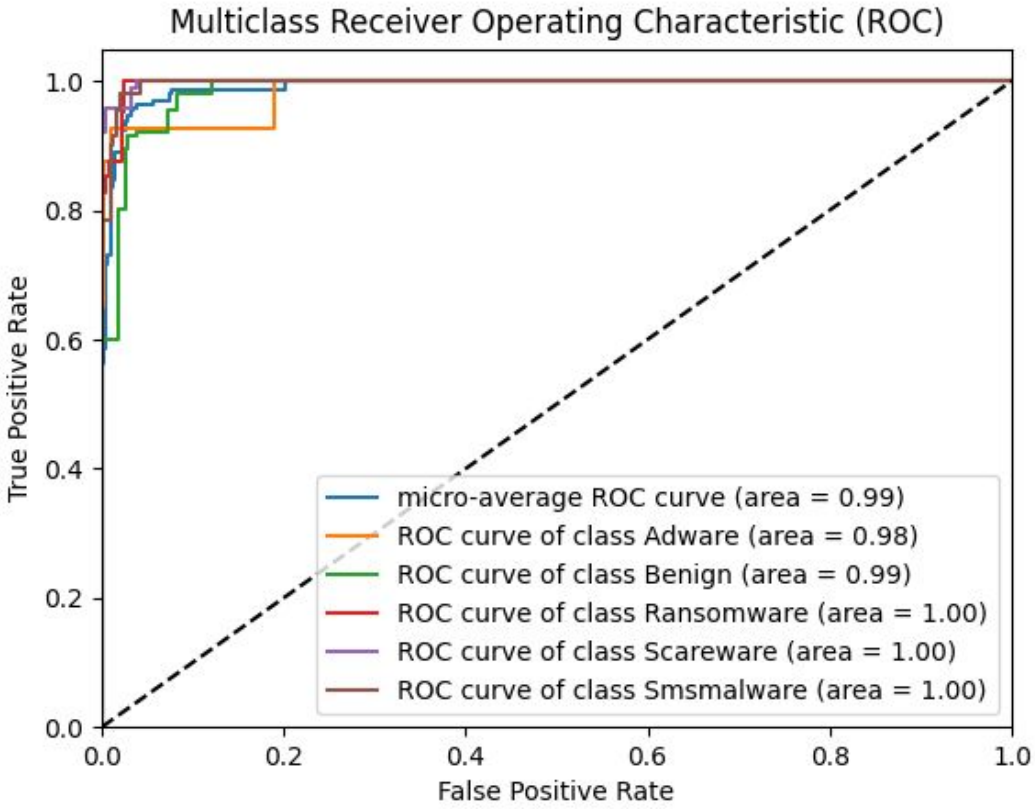
Florestas Aleatórias

Classificador | Florestas Aleatórias



Pasta 4

Classificador | Florestas Aleatórias



Pasta 4

Classificador | Florestas Aleatórias

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Precisão	1	0.9677	0.8421	0.9769	0.9864	0.974
	2	0.9747	0.8077	0.96	**1.0**	0.9932
	3	0.9548	**0.7853**	0.9231	0.986	0.979
	4	0.9551	**0.7857**	0.8942	0.986	0.9923
	5	0.9419	0.7944	0.9494	**1.0**	0.9796
Recall	1	0.9375	0.9474	**1.0**	0.8788	0.974
	2	0.9565	0.9735	**1.0**	**0.8253**	0.9545
	3	0.9193	0.9205	**1.0**	0.8494	0.9091
	4	0.9255	0.9408	**1.0**	0.8545	0.8431
	5	0.9068	0.9408	**1.0**	0.8485	0.9412

Classificador | Florestas Aleatórias

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
F1-Score	1	0.9524	0.8916	0.9883	0.9295	0.974
	2	0.9655	0.8829	0.9796	0.9043	0.9735
	3	0.9367	0.8476	0.96	0.9126	0.9428
	4	0.9401	0.8563	0.9441	0.9156	0.9117
	5	0.9241	0.8614	0.9741	0.918	0.96
AUC	1	0.9884	0.988	0.9997	0.9978	0.999
	2	0.9882	0.9915	0.9999	0.9966	0.9982
	3	0.9788	0.9817	0.9969	0.9977	0.9973
	4	0.985	0.9861	0.9969	0.9984	0.9967
	5	0.9776	0.9798	0.9995	0.9984	0.9995

Classificador | Florestas Aleatórias

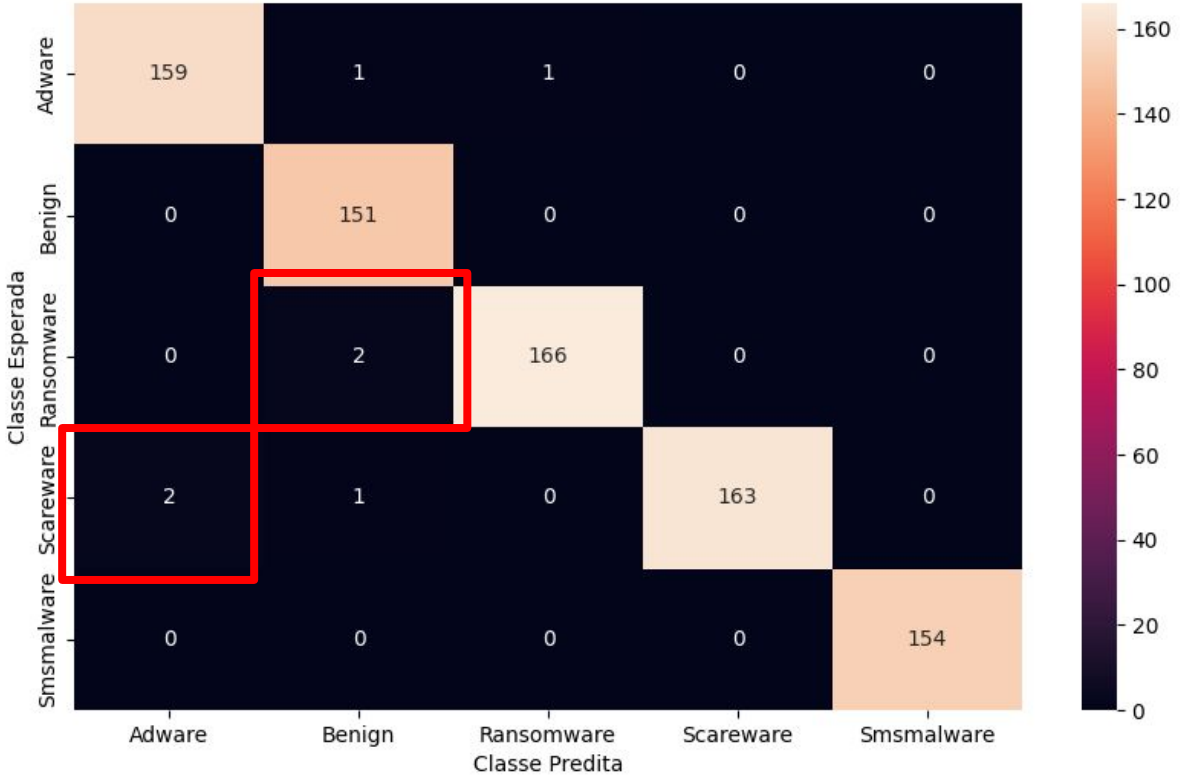
Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Acurácia	1	**0.9475**				
	2	0.9413				
	3	0.92				
	4	**0.9137**				
	5	0.9275				

Classificador | Florestas Aleatórias

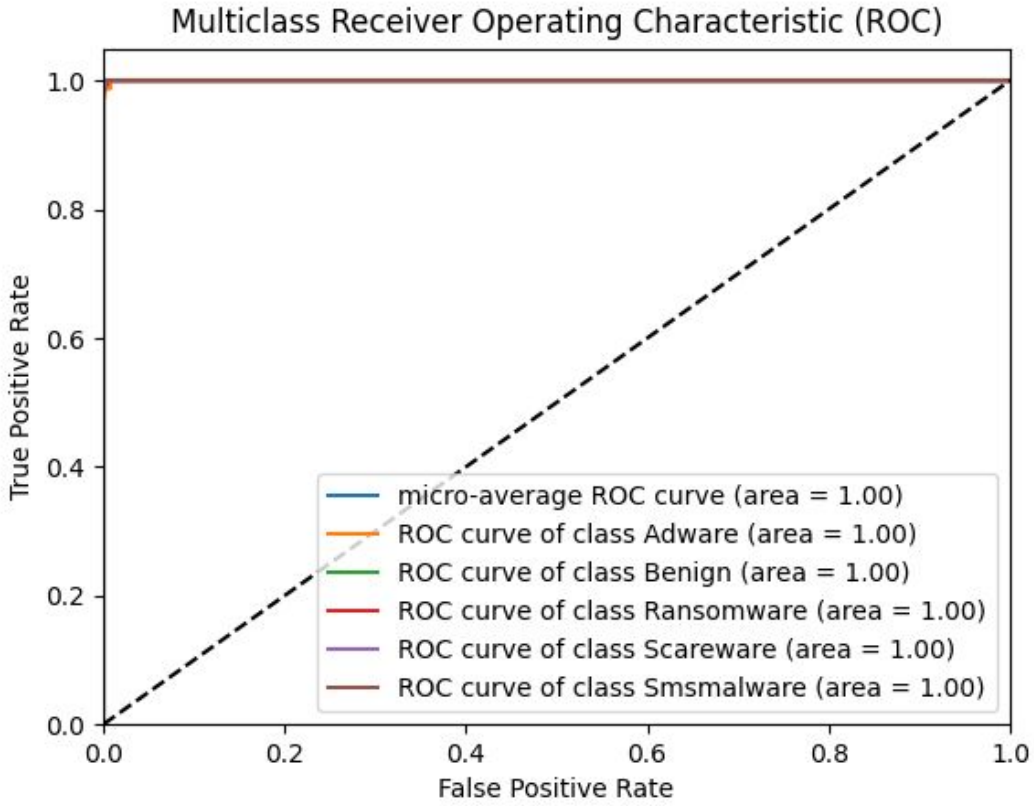
- **Considerações:**
  - **Incerteza do modelo:**
    - Prevê Benigno quando Adware ou Scareware
    - Prevê Ransomware quando SMS Malware
    - Prevê Scareware ou SMS Malware quando Benigno
  - **Precisão** da classe Benigna afetada pelos Falsos Positivos (78%)
  - **Recall** acima de 82% → redução de Falsos Negativos
  - **Hiperparâmetros:**
    - Poucas árvores e não tão profundas
    - Treinamento rápido (menos de 1 segundo)
    - **Melhoria:** aumentar os intervalos da *Grid Search*

KNN

Classificador | KNN



Classificador | KNN



Classificador | KNN

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Precisão	1	0.9938	1.0	1.0	1.0	0.9872
	2	0.9877	0.9869	1.0	1.0	1.0
	3	0.9876	**0.9742**	0.994	1.0	1.0
	4	1.0	0.9935	0.9941	1.0	0.9935
	5	1.0	1.0	0.9941	0.988	1.0
Recall	1	1.0	**1.0**	**1.0**	0.9818	1.0
	2	1.0	**1.0**	**1.0**	**0.9759**	1.0
	3	0.9876	**1.0**	0.9881	0.9819	1.0
	4	0.9938	**1.0**	**1.0**	0.9939	0.9935
	5	0.9938	**1.0**	**1.0**	1.0	0.9869

Classificador | KNN

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
F1-Score	1	0.9969	1.0	1.0	0.9908	0.9935
	2	0.9938	0.9934	1.0	0.9878	1.0
	3	0.9876	0.9869	0.991	0.9909	1.0
	4	0.9969	0.9967	0.9971	0.997	0.9935
	5	0.9969	1.0	0.9971	0.994	0.9934
AUC	1	**1.0**	**1.0**	**1.0**	0.9999	**1.0**
	2	**1.0**	**1.0**	**1.0**	0.9969	**1.0**
	3	0.9999	**1.0**	**1.0**	**1.0**	**1.0**
	4	**1.0**	**1.0**	0.9992	**1.0**	0.9967
	5	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**

Classificador | KNN

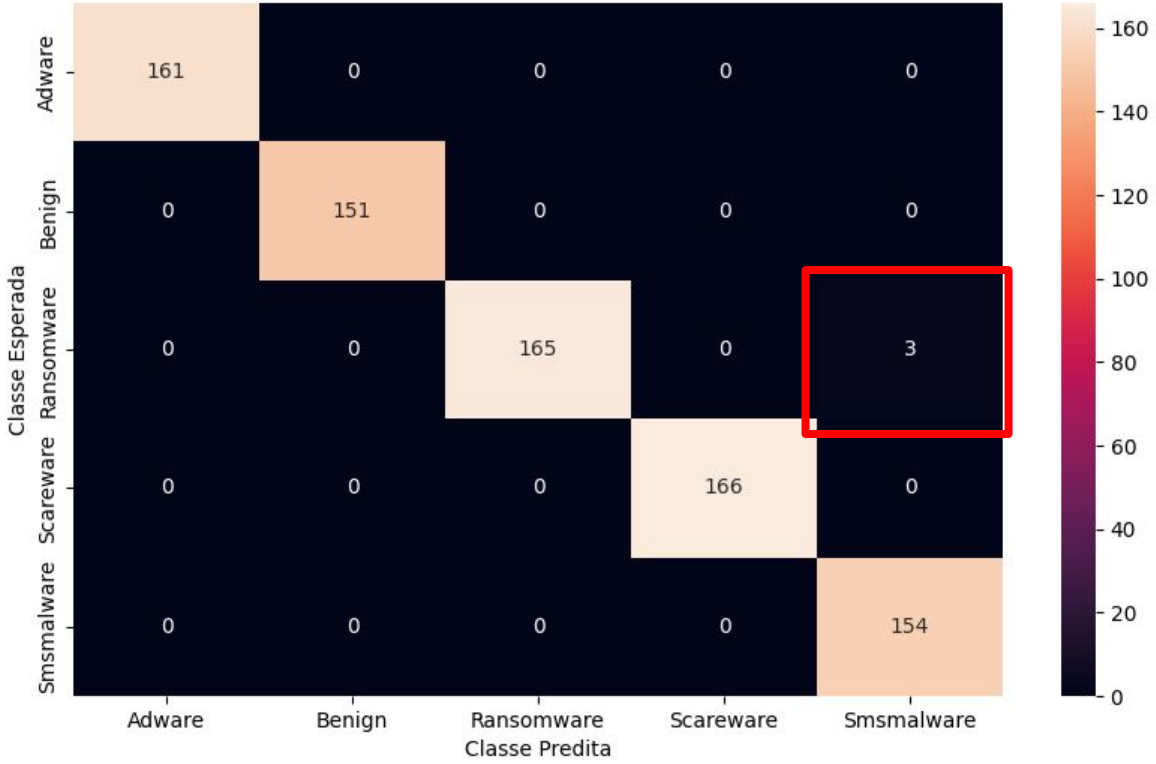
Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Acurácia	1	0.9962				
	2	**0.995**				
	3	**0.9912**				
	4	0.9962				
	5	0.9962				

Classificador | KNN

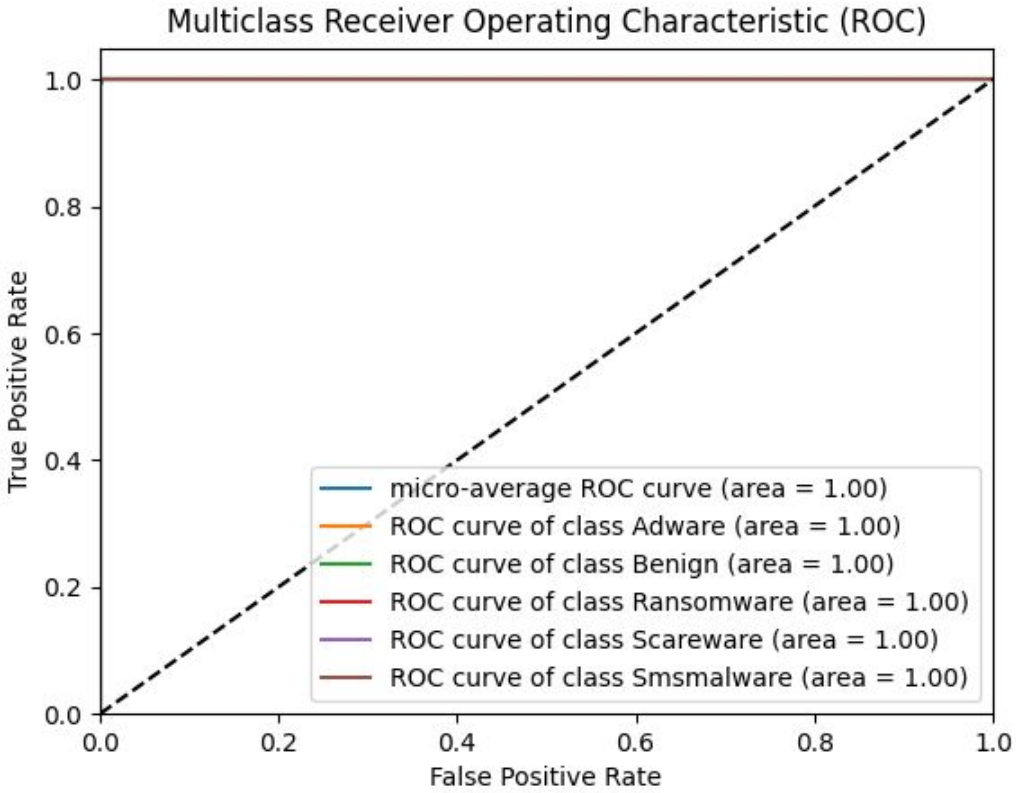
- **Considerações:**
  - **Pouca incerteza do modelo:**
    - Prevê SMS Malware, Adware ou Benigno quando Scareware
    - Prevê Ransomware ou Scareware quando SMS Malware
  - **Curvas ROC** próximas à curva ideal (AUC de 1,0)
  - **Recall** próximo à 1 em todas as pastas ➔ redução de Falsos Negativos
  - **Hiperparâmetros:**
    - 5 vizinhos de mesmo peso com a distância sendo a Euclidiana
    - Bons vetores de características
    - Amostras da mesma classe próximas no espaço
    - Amostras de classes diferentes distantes
    - Intervalos de otimização escolhidos se encaixaram bem ao problema

Gradient Boosting

Classificador | Gradient Boosting



Classificador | Gradient Boosting



Classificador | Gradient Boosting

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Precisão	1	1.0	1.0	1.0	1.0	1.0
	2	1.0	1.0	1.0	1.0	1.0
	3	1.0	1.0	1.0	1.0	**0.9809**
	4	1.0	1.0	1.0	1.0	1.0
	5	1.0	1.0	1.0	1.0	1.0
Recall	1	1.0	1.0	1.0	1.0	1.0
	2	1.0	1.0	1.0	1.0	1.0
	3	1.0	1.0	**0.9821**	1.0	1.0
	4	1.0	1.0	1.0	1.0	1.0
	5	1.0	1.0	1.0	1.0	1.0

Classificador | Gradient Boosting

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
F1-Score	1	1.0	1.0	1.0	1.0	1.0
	2	1.0	1.0	1.0	1.0	1.0
	3	1.0	1.0	**0.991**	1.0	**0.9904**
	4	1.0	1.0	1.0	1.0	1.0
	5	1.0	1.0	1.0	1.0	1.0
AUC	1	1.0	1.0	1.0	1.0	1.0
	2	1.0	1.0	1.0	1.0	1.0
	3	1.0	1.0	1.0	1.0	1.0
	4	1.0	1.0	1.0	1.0	1.0
	5	1.0	1.0	1.0	1.0	1.0

Classificador | Gradient Boosting

Métrica	Fold	Classes				
		Adware	Benign	Ransomware	Scareware	SMS Malware
Acurácia	1	1.0				
	2	1.0				
	3	0.9962				
	4	1.0				
	5	1.0				

Classificador | Gradient Boosting

- **Considerações:**
  - **Pouca incerteza do modelo:**
    - Confusão entre a classe SMS Malware com a Ransomware
  - A confusão do modelo teve repercussões na **Precisão**, **Recall** e **F1-Score**
  - **Acurácia** só diminuiu em uma das pastas (99,62%)
  - **Hiperparâmetros:**
    - Melhores valores não muito distantes do padrão
    - Número baixo de estimadores e pouca profundidade
    - Treinamento rápido ➔ GB foi eficiente para solucionar o problema
  - **Esperado:**
    - Objetivo de conseguir resultados melhores do que o KNN ➔ alcançado

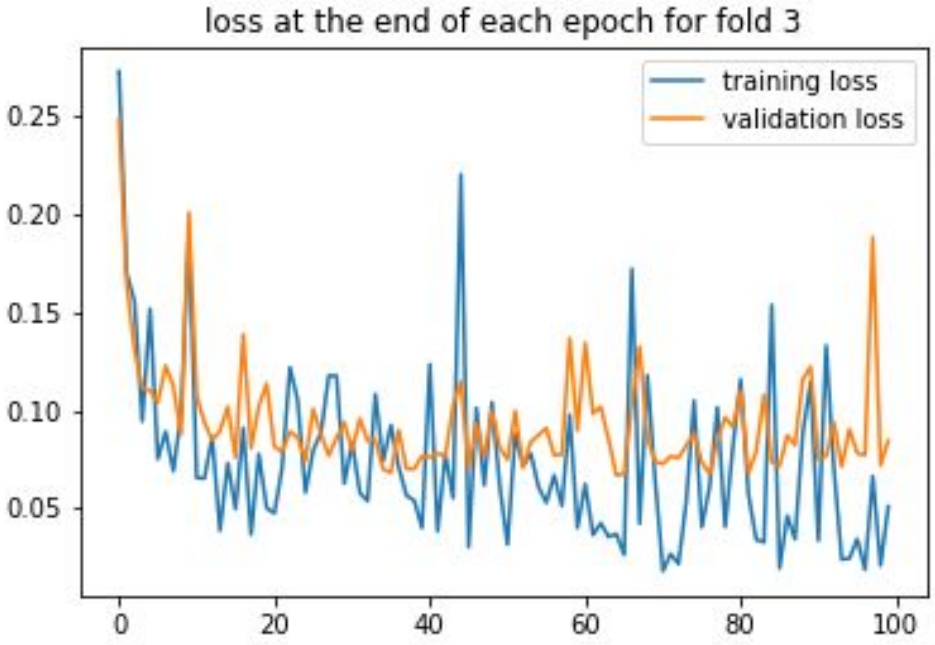
Modelos Baseados em Aprendizado Profundo

Classificador | Modelo Linear

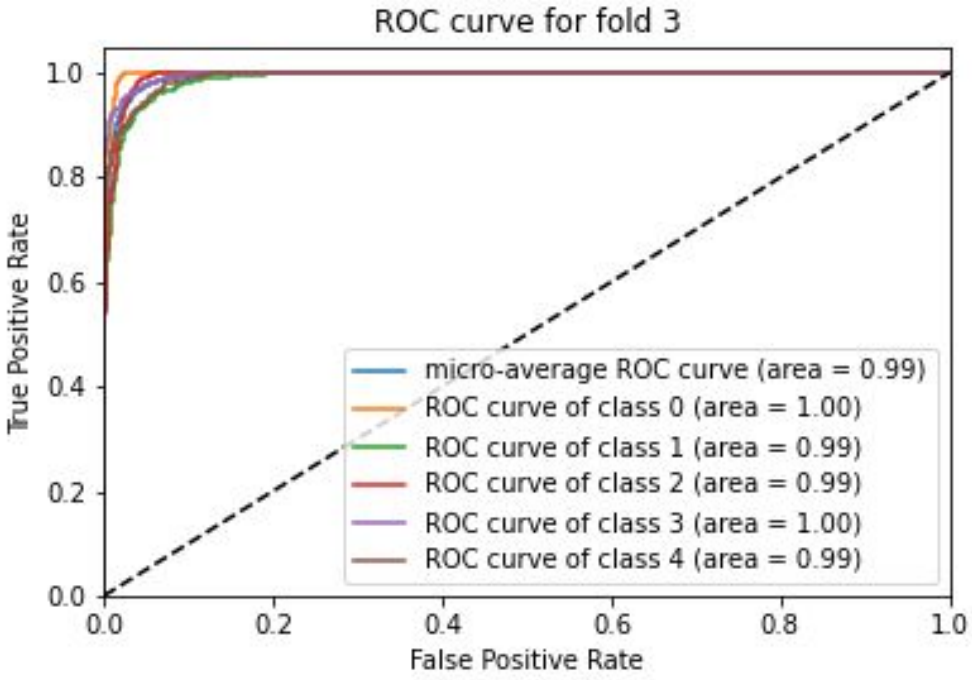
- Composta por camadas lineares.
- Este tipo de camada usa uma operação linear, ou seja, a saída de cada neurônio é formada em função de suas entradas.
- Função de ativação *ReLU.*

Linear(130, 512)
ReLU ( )
Linear(512, 1024)
ReLU ( )
Linear(1024, 4096)
ReLU ( )
Linear(4096, 4096)
ReLU ( )
Linear(4096, 2048)
ReLU ( )
Linear(2048, 1024)
ReLU ( )
Linear(1024, 5)

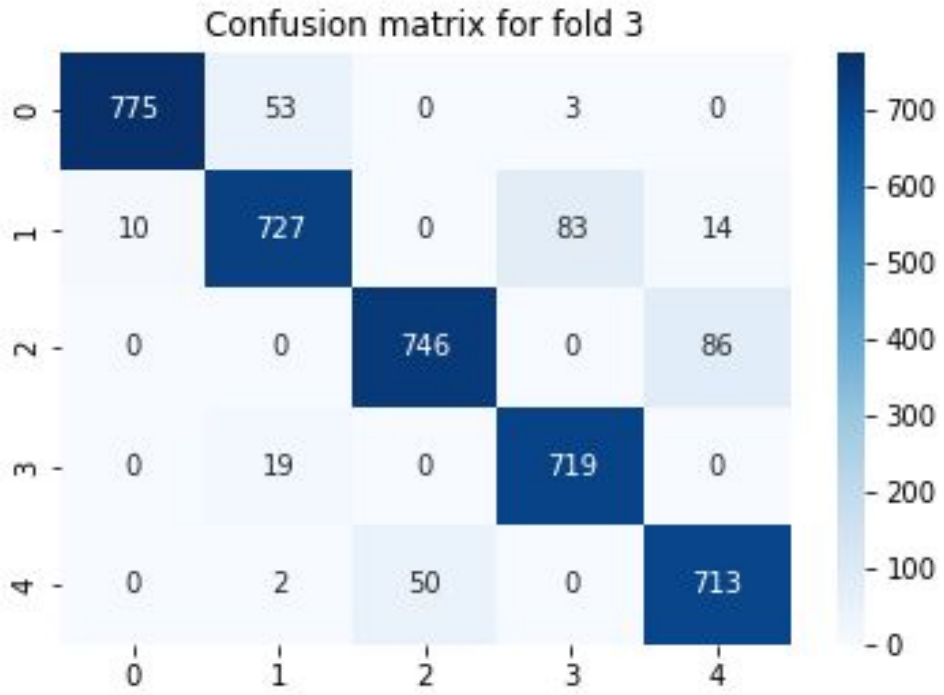
Classificador | Modelo Linear



Classificador | Modelo Linear



Classificador | Modelo Linear



Classificador | Modelo Linear

Métrica	Fold	Classes				
		0	1	2	3	4
Precisão	0	0.865460	0.895070	0.918310	0.969251	0.903346
	1	0.898246	0.918845	0.905830	0.974392	0.978523
	2	0.912633	0.911444	0.905896	0.978426	0.934579
	3	0.932611	0.871703	0.896635	0.974255	0.932026
	4	0.946996	0.923845	**0.833162**	0.973440	0.995208
Recall	0	0.979616	**0.824214**	0.898072	0.921220	0.918136
	1	0.993532	0.860825	0.985366	0.931457	0.895577
	2	0.982211	0.880263	0.944444	0.941392	0.888325
	3	0.987261	0.907615	0.937186	0.893168	0.876999
	4	0.979294	0.920398	0.996305	0.949482	**0.787611**

Classificador | Modelo Linear

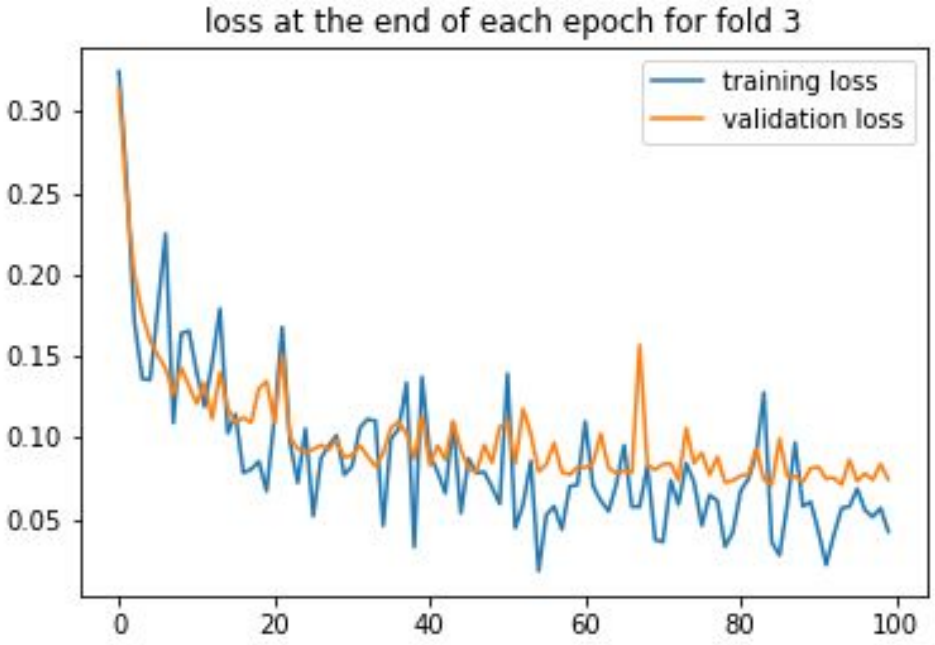
Métrica	Fold	Classes				
		0	1	2	3	4
F1-Score	0	0.919010	**0.858182**	0.908078	0.944625	0.910681
	1	0.943489	0.888889	0.943925	0.952441	0.935215
	2	0.946144	0.895582	0.924769	0.959552	0.910865
	3	0.959158	0.889297	0.916462	0.931951	0.903676
	4	0.962874	0.922118	0.907459	0.961311	0.879323

Classificador | Modelo Convolucional

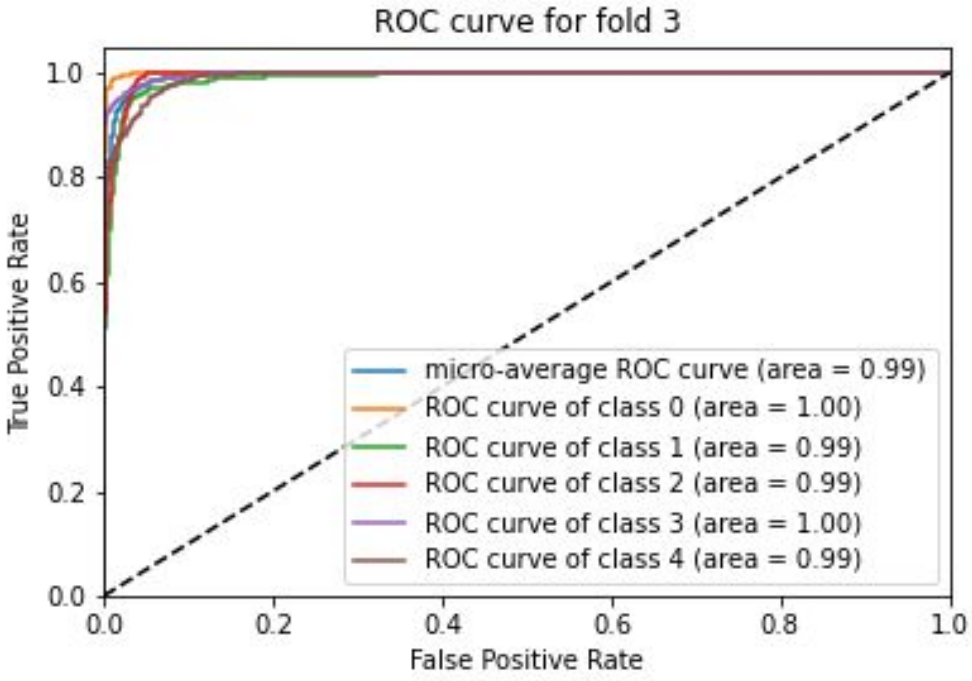
- A entrada foi interpretada como tensores de 130 valores e 1 canal.
- Operação *MaxPool* foi utilizada de maneira a ressaltar valores mais significativos.
- Aplica-se a operação *Flatten* para linearizar o resultado pelas camadas densas (lineares).

Conv1d(1, 64, kernel=3, stride=1)
ReLU ( )
Conv1d(64, 64, kernel=3, stride=1)
ReLU ( )
MaxPool1d(kernel=3, stride=1)
Conv1d(64, 256, kernel=3, stride=1)
ReLU ( )
Conv1d(256, 256, kernel=3, stride=1)
ReLU ( )
MaxPool1d(kernel=3, stride=1)
Flatten ( )
Linear(30208, 2048)
ReLU ( )
Linear(2048, 1024)
ReLU ( )
Linear(1024, 5)

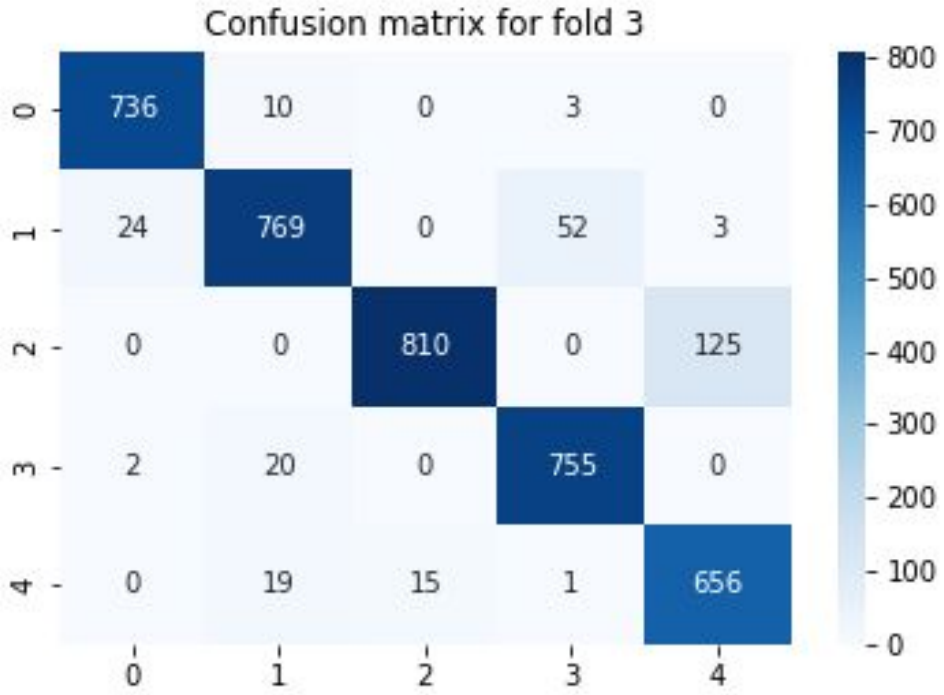
Classificador | Modelo Convolucional



Classificador | Modelo Convolucional



Classificador | Modelo Convolucional



Classificador | Modelo Convolucional

Métrica	Fold	Classes				
		0	1	2	3	4
Precisão	0	0.987585	0.939086	0.896595	0.992136	0.914286
	1	0.971985	0.935301	0.867059	0.990515	0.957746
	2	0.888889	0.925982	0.868282	0.959410	0.961593
	3	0.982644	0.906840	**0.866310**	0.971686	0.949349
	4	0.920732	0.940026	0.912609	0.977749	0.880896
Recall	0	0.988701	0.964798	0.933071	0.948622 .0	0.893401
	1	0.987624	0.957027	0.981358	0.938383	0.848939
	2	0.991060	**0.819519**	0.980952	0.948905	**0.837670**
	3	0.965879	0.940098	0.981818	0.930949	**0.836735**
	4	0.990814	0.894541	0.889294	0.945570	0.910976

Classificador | Modelo Convolucional

Métrica	Fold	Classes				
		0	1	2	3	4
F1-Score	0	0.919010	**0.858182**	0.908078	0.944625	0.910681
	1	0.943489	0.888889	0.943925	0.952441	0.935215
	2	0.946144	0.895582	0.924769	0.959552	0.910865
	3	0.959158	0.889297	0.916462	0.931951	0.903676
	4	0.962874	0.922118	0.907459	0.961311	0.879323

Considerações Finais

- A seleção de Características contribuiu para os bons resultados, reduzindo a quantidade de atributos de 481 para 50.
  - Treinamento dos modelos ocorreu mais rápido.
  - Considerando apenas as informações mais relevantes para a efetuar a classificação.

- Métodos baseados em Aprendizado de Máquina mostraram-se superiores na classificação dos malwares presentes no *Dataset*.

- Em futuros trabalhos, pretende-se estudar os efeitos de outros hiperparâmetros e arquiteturas de redes neurais.