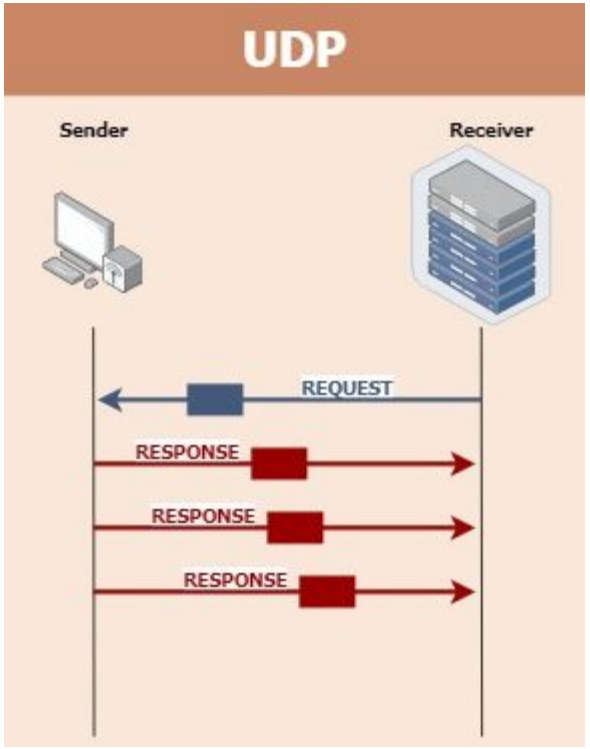
Detecção de protocolos de aplicação

67_UDP_DhcPs e 137_UDP

Introdução

- UDP (sigla para User Datagram Protocol)
- Protocolo da camada de transporte
- Simples
- Não requer handshake
- Bastante vulnerável



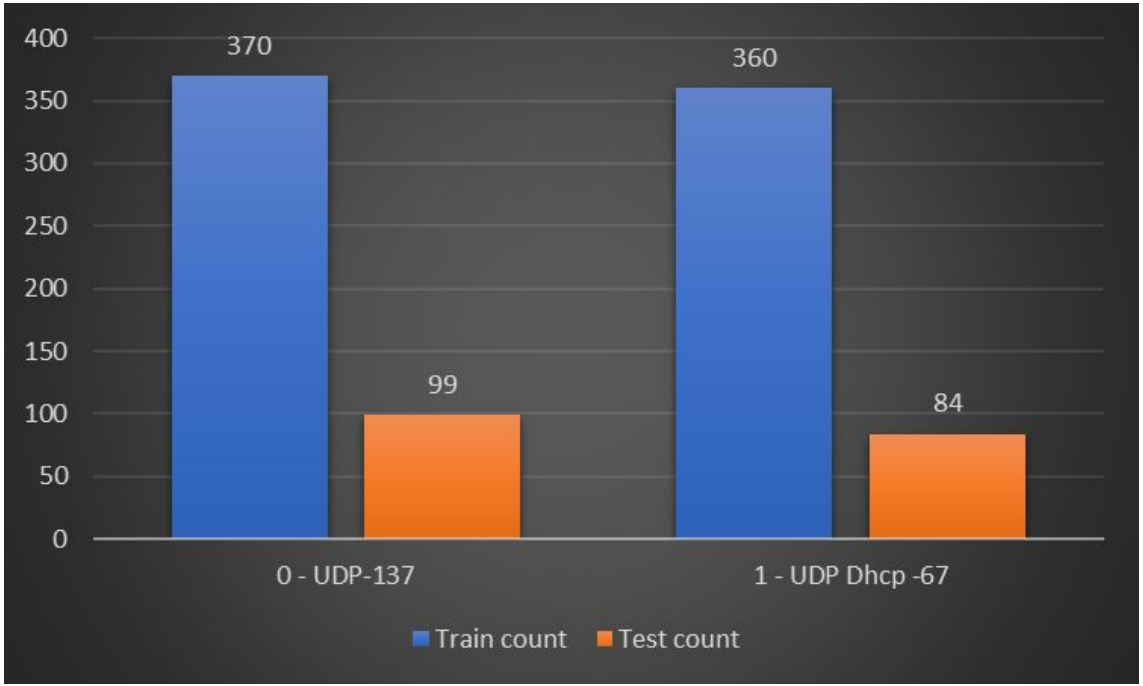
Dataset

- https://www.inf.ufpr.br/gregio/CI1030/final/appIdent.json
- Analisado e “Limpo”
- Juntado de forma a tem os mesmos parâmetros para comparação
- Dividido em : Treino(80%) e teste (20%)

index	
ApplicationProtocolName	object
ApplicationProtocolNameFull	object
MinInterArrivalTimePacketsUpAndDownFlow.FeatureValue	float64
MinInterArrivalTimePacketsUpAndDownFlow.ModelValues	float64
MinInterArrivalTimePacketsUpAndDownFlow.Weight	float64
MinInterArrivalTimePacketsUpAndDownFlow.NormalizedWeight	float64
MinInterArrivalTimePacketsUpAndDownFlow.Mean	float64
MinInterArrivalTimePacketsUpAndDownFlow.StdDev	float64
MinInterArrivalTimePacketsUpAndDownFlow.Max	float64
MinInterArrivalTimePacketsUpAndDownFlow.Min	float64
MinInterArrivalTimePacketsUpFlow.FeatureValue	float64
MinInterArrivalTimePacketsUpFlow.ModelValues	float64
MinInterArrivalTimePacketsUpFlow.Weight	float64
MinInterArrivalTimePacketsUpFlow.NormalizedWeight	float64
MinInterArrivalTimePacketsUpFlow.Mean	float64
MinInterArrivalTimePacketsUpFlow.StdDev	float64
MinInterArrivalTimePacketsUpFlow.Max	float64
MinInterArrivalTimePacketsUpFlow.Min	float64
MinInterArrivalTimePacketsDownFlow.FeatureValue	float64
MinInterArrivalTimePacketsDownFlow.ModelValues	float64
MinInterArrivalTimePacketsDownFlow.Weight	float64
MinInterArrivalTimePacketsDownFlow.NormalizedWeight	float64
MinInterArrivalTimePacketsDownFlow.Mean	float64
MinInterArrivalTimePacketsDownFlow.StdDev	float64
MinInterArrivalTimePacketsDownFlow.Max	float64

Dataset

A amostragem da divisão dos dados ficou em:



Metodologia

Classificação dos protocolos UDP Dhcp -67 e UDP-13 utilizando o Dataset rotulado para treinar os algoritmos K-Nearest Neighbors (KNN), Random Forest e Multi Layer Perceptron (MLP), com implementação na biblioteca sklearn. Para a configuração dos parâmetros dos classificadores foi utilizado o algoritmo Grid Seach para predizer os melhores parâmetros de cada um deles.

Implementação

Colaboratory ou "Colab" permite escrever código Python no seu navegador, com:

- Nenhuma configuração necessária
- Acesso gratuito a GPUs
- Compartilhamento fácil
- Você pode ser um estudante,

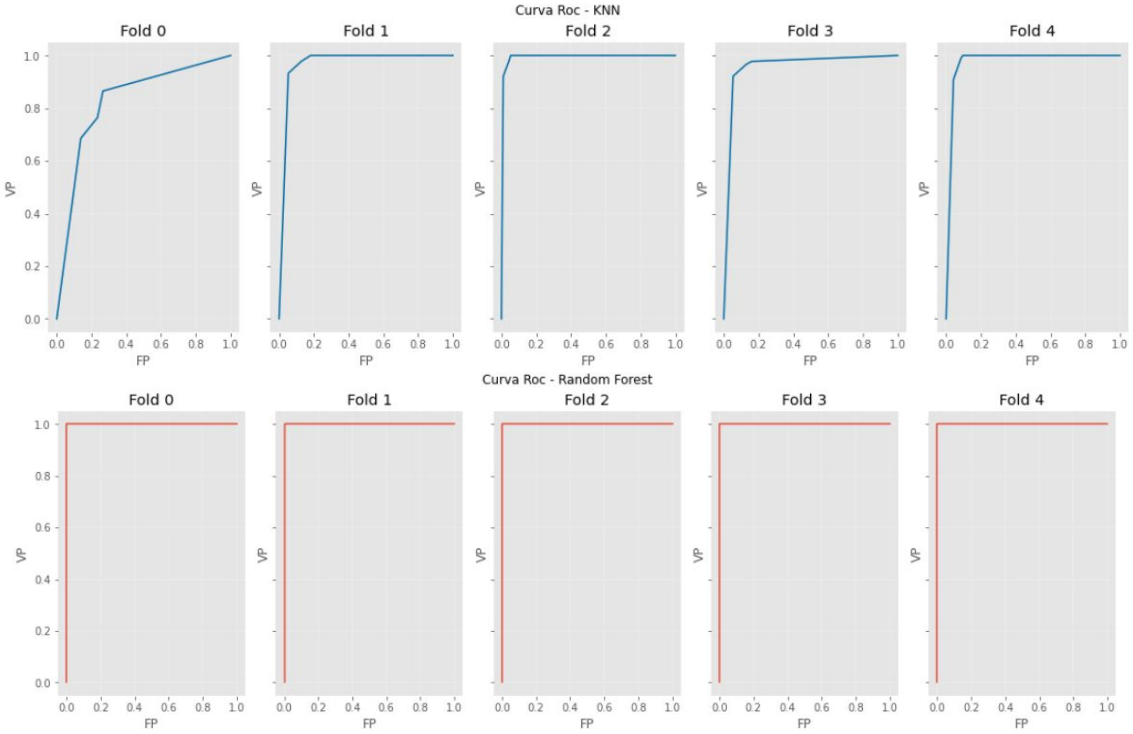
Código:

https://colab.research.google.com/drive/10-50NKxXO8Caks67rPQErNZCSdV9S5r2#scrollTo=04g1jitfVoia

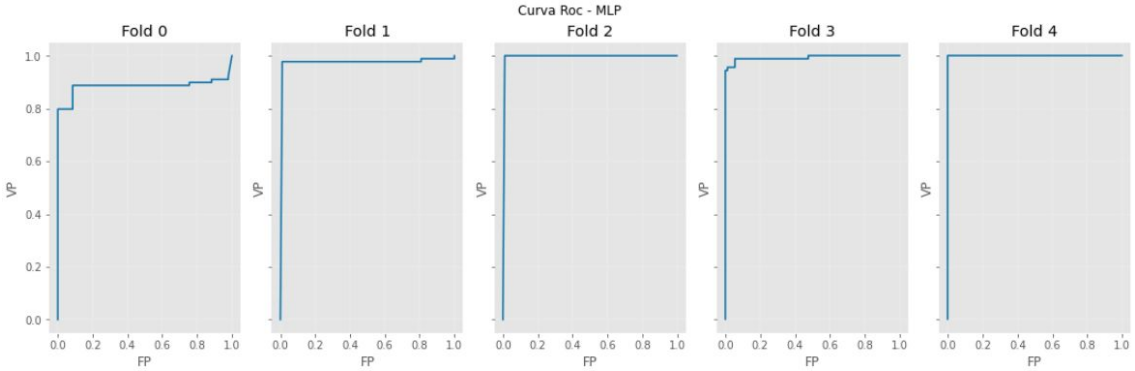
Resultados:

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
KNN					
Matriz de confusão	72 22 21 68	82 12 2 87	90 4 2 87	81 12 3 86	86 8 1 87
Precisão	0.77	0.88	0.92	0.88	0.92
Recall	0.76	0.98	1.00	0.97	0.99
f1score	0.76	0.93	0.96	0.92	0.95
Random Forest					
Matriz de confusão	94 0 0 89	94 0 0 89	94 0 0 89	94 0 0 89	94 0 0 89
Precisão	1.00	1.00	1.00	1.00	1.00
Recall	1.00	1.00	1.00	1.00	1.00
f1score	1.00	1.00	1.00	1.00	1.00
MLP					
Matriz de confusão	78 16 10 79	91 3 2 87	86 8 0 89	92 1 4 85	86 8 1 87
Precisão	0.83	0.97	0.92	0.99	1.00
Recall	0.89	0.98	1.00	0.96	1.00
f1score	0.86	0.97	0.96	0.97	1.00

Resultados:


Curva Roc - KNN
Fold 0
Fold 1
Fold 2
Fold 3
Fold 4
VP
FP
Curva Roc - Random Forest
Fold 0
Fold 1
Fold 2
Fold 3
Fold 4
VP
FP

Resultados:


Curva Roc - MLP
Fold 0
Fold 1
Fold 2
Fold 3
Fold 4
VP
FP
0.0
0.2
0.4
0.6
0.8
1.0

Reprodutibilidade

- Datasets: https://drive.google.com/drive/folders/1FcRpscUYHNkOgeqSfk4kWRpRGulHxCtz?usp=sharing
- Implementação: https://colab.research.google.com/drive/10-50NKxXO8Caks67rPQErNZCSdV9S5r2?usp=sharing
- Link para vídeo e apresentação: https://github.com/ACBozzi/Ci-nciaDados/tree/main
-

Referências

UFSM. Introdução a Mineração de textos com Python. 2021. Disponível em: <https://www.ufsm.br/pet/sistemas-de-informacao/2021/07/12/introducao-a-mineracao-de-textos-com-python/>. Acesso em: 15 nov. 2021.

POSTEL, J. RFC 768 - User Datagram Protocol. Internet Engineering Task Force (IETF), p.3. 1980.

Ferramenta IPERF: geração e medição de Tráfego TCP e UDP IPERF tool: generation and evaluation of TCP and UDP data traffic. 2014. Disponível em <http://revistas.cbpf.br/index.php/nt/article/view/75/67> Acesso em: 16 dez. 2021.