IWSHAP: Um Método de Seleção Incremental de Características para Redes CAN baseado em Inteligência Artificial Explicável (XAI)

Felipe H. Scherer, Felipe N. Dresch
Silvio E. Quincozes, Diego Kreutz,
Vagner E. Quincozes

LEA
Laboratório de Estudos Avançados em Computação

DEFESA Since 2022 DC CIBERNÉTICA

Como a inteligência artificial explicável (XAI) pode contribuir para uma seleção de características mais rápida e precisa?

Contexto

- Redes *Controller Area Network* (CAN)
- Troca de informação entre as Unidades de Controle Eletrônico (ECU)
- Limitações Computacionais e Vulnerabilidades

Motivação

Fonte: cbc

Ciência

Hackers destroem motor de Jeep em movimento em rodovia em demonstração de segurança

Fiat Chrysler diz que patch foi lançado para a vulnerabilidade mais séria envolvida

Thomson Reuters · Postado: 22 de julho de 2015 9:16 AM EDT | Última atualização: 23 de julho de 2015

Motivação

Fonte: The Guardian

Equipe de hackers assume controle remoto do Tesla Model S a 12 milhas de distância

Pesquisadores chineses conseguiram interferir nos freios do carro, travas das portas e outros recursos eletrônicos, demonstrando um ataque que poderia causar estragos

Motivação

- Emersão de Sistema de Detecção de Intrusão (IDSs) em Redes CAN
- Falta de explicabilidade
- Inteligência Artificial Explicável

Fundamentação

- Seleção eficaz de características:
  - Elevar o desempenho dos IDSs!
- Seleção por:
  - Filtragem
  - *Wrapper*
  - *Embedded*
  - *Abordagens Híbridas: Filtragem + Wrapping*

Problema(s)

- Há uma lacuna na literatura!
  - Trabalhos atuais tratam da seleção de maneira ineficiente!

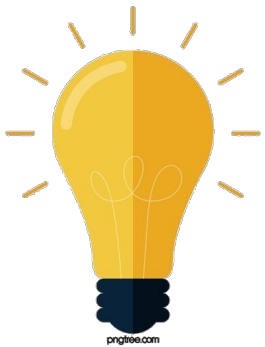
Desafio(s)

- Como **maximizar a assertividade** de um IDS enquanto ganha **desempenho computacional**?

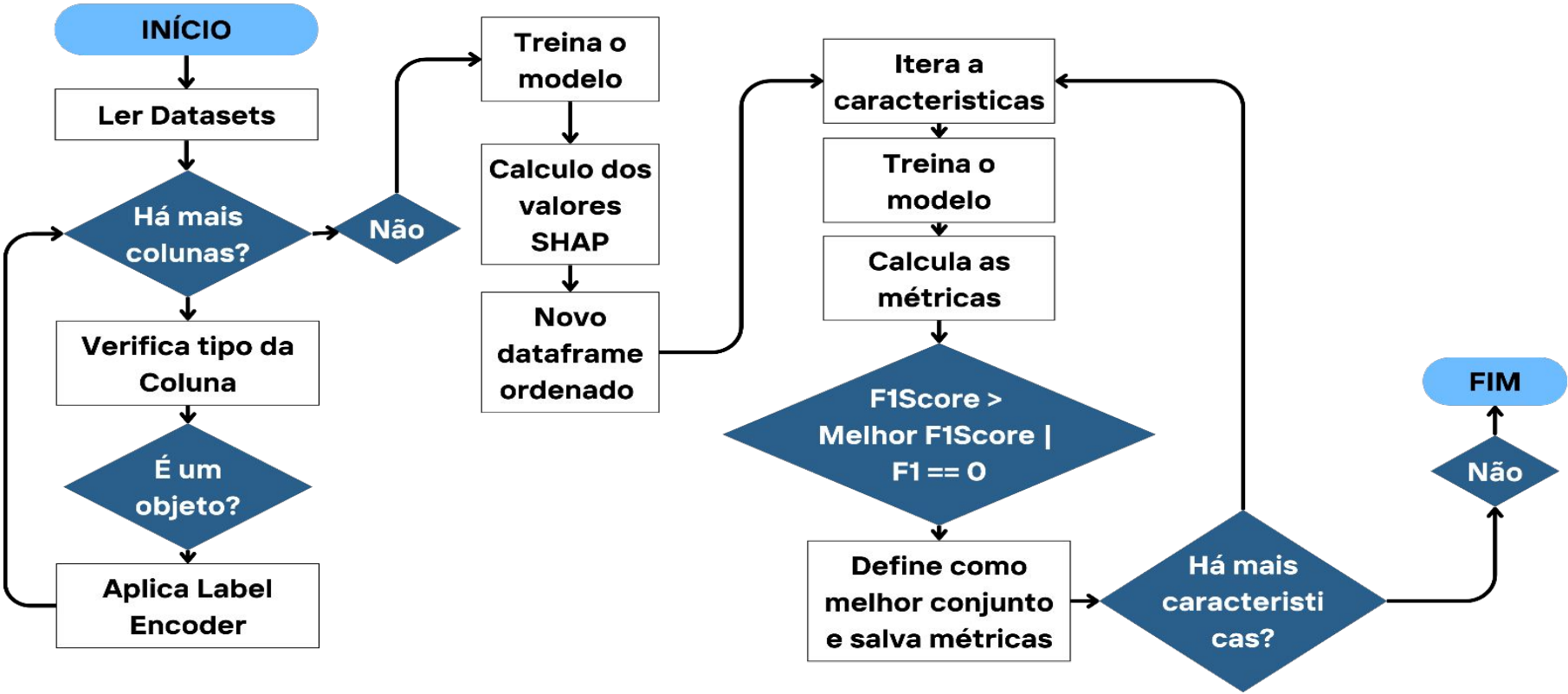
Precisamos processar menos informações, mas sem perder informações relevantes!

Proposta

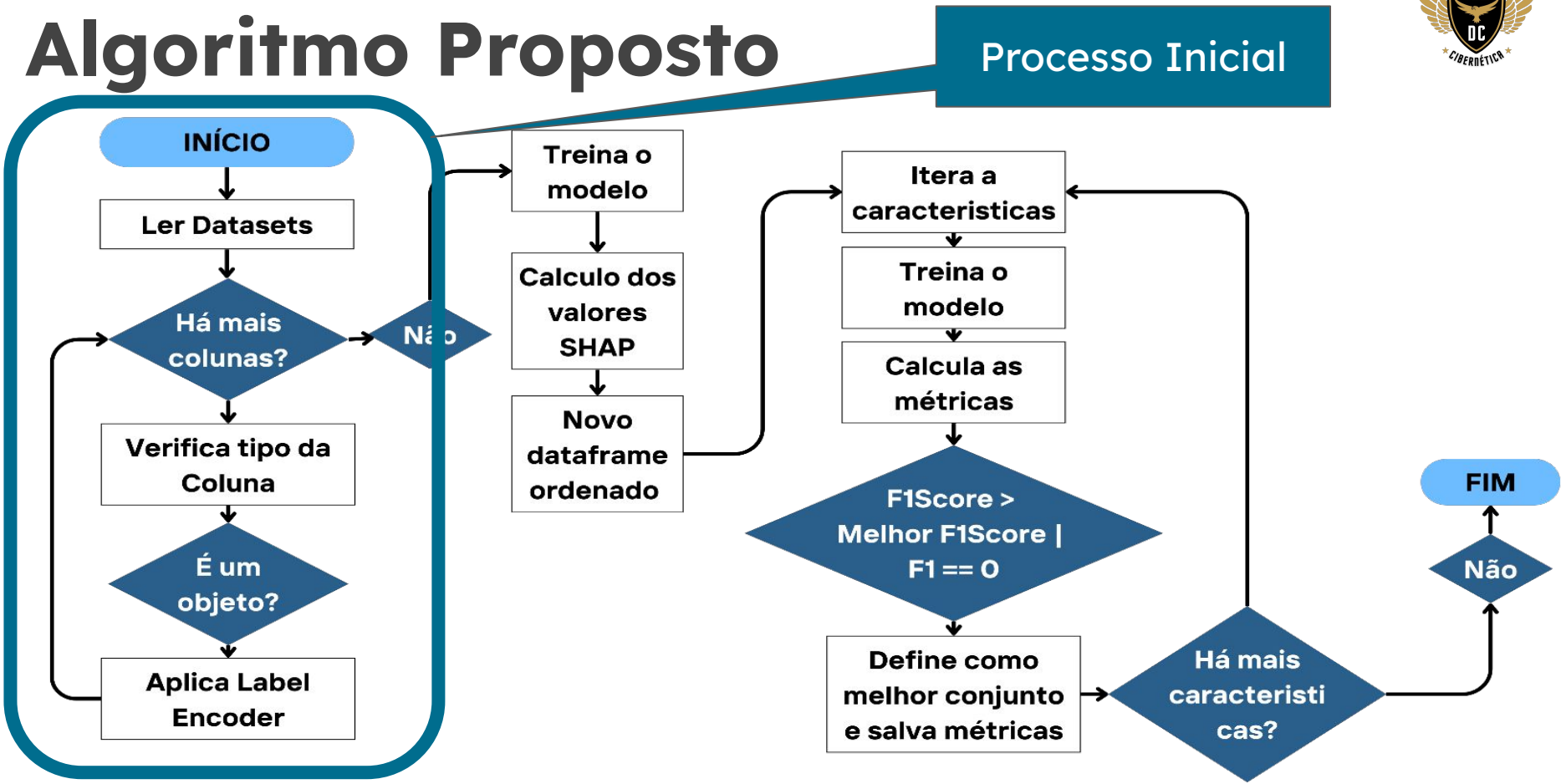
- **Combinar:**
  - Incremental Wrapper Subset Selection (IWSS)
  - Inteligência artificial explicável (XAI)!



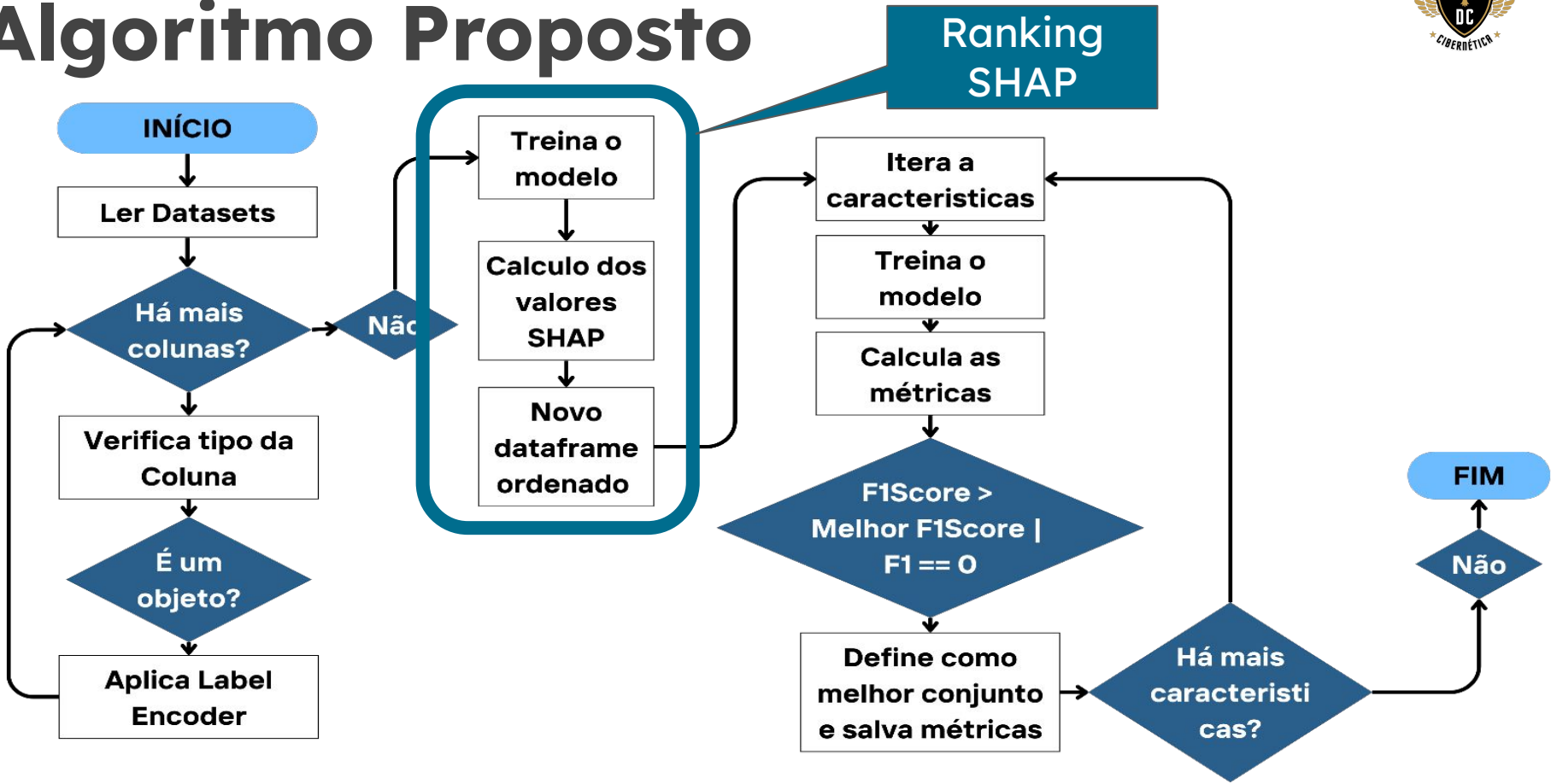
Algoritmo Proposto



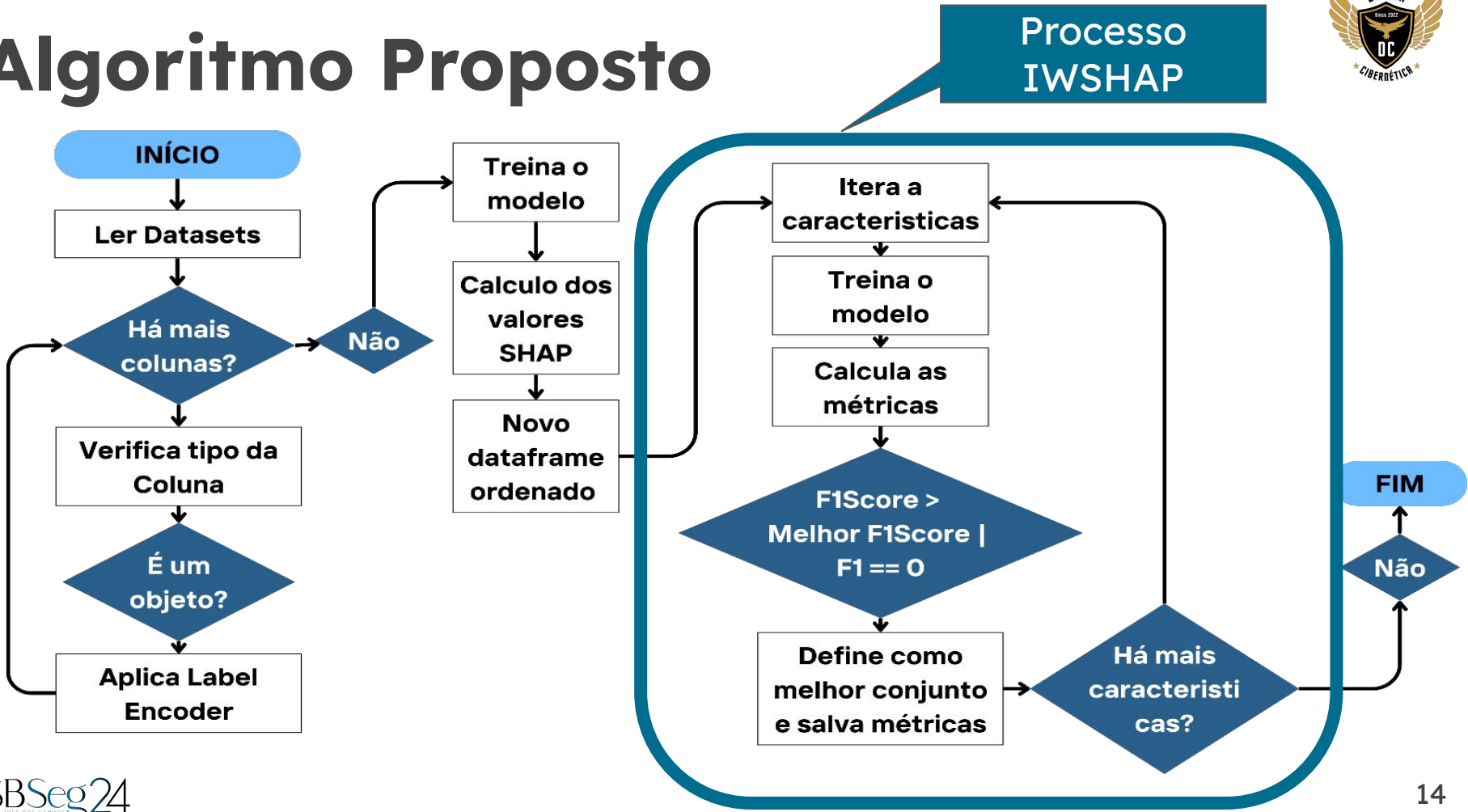
Algoritmo Proposto



Algoritmo Proposto



Algoritmo Proposto



Cenário de Experimentação

- Hardware de experimentos:
  - Ubuntu / AMD Ryzen 7 5800x 8-Core
  - 64 GB RAM
- Bibliotecas utilizadas:

Cenário de Experimentação

- Dados:
  - Dataset **X-CANIDS**
  - Ataque de **suspensão**!
    - Foco no AID 2B0h

Avaliação

- Métricas de avaliação:
  - Quantia de características
  - Tempo de execução
  - Métricas de desempenho do modelo

Avaliação

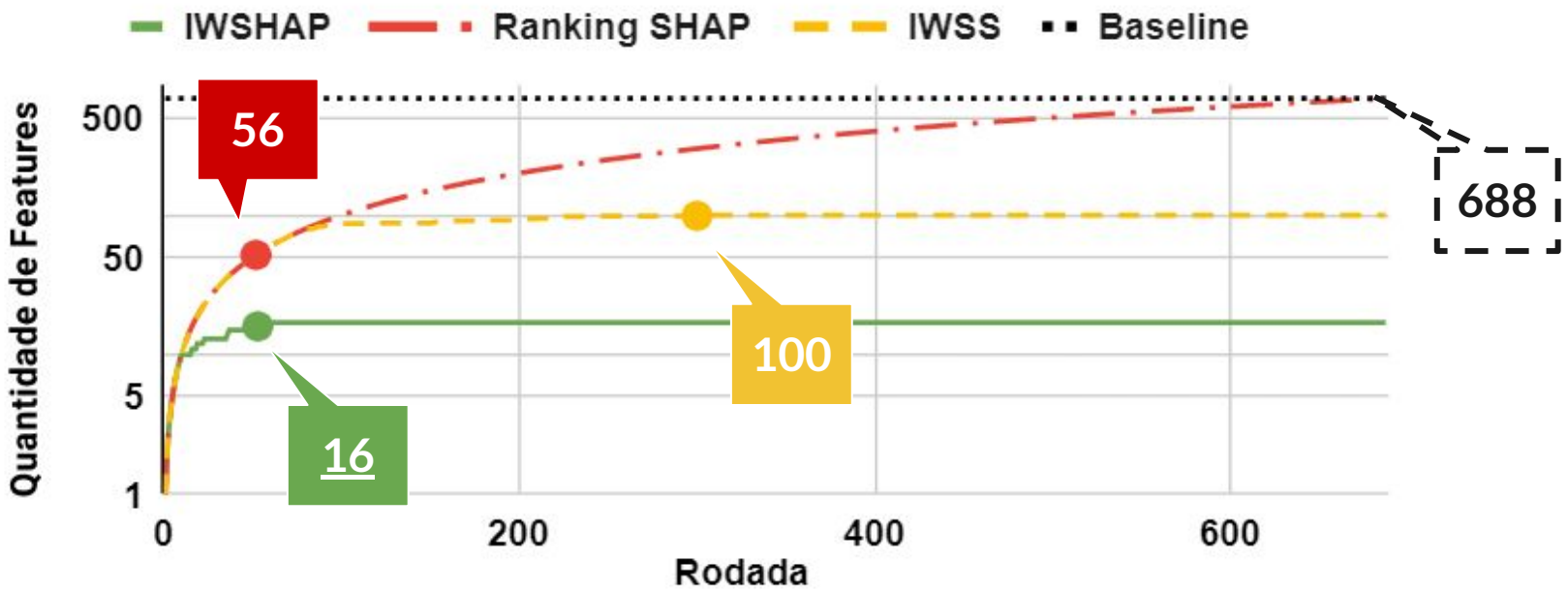
- Métodos comparativos:
  - IWSS
  - Ranking SHAP
  - Baseline

Resultados

- IWSHAP superou o desempenho dos demais

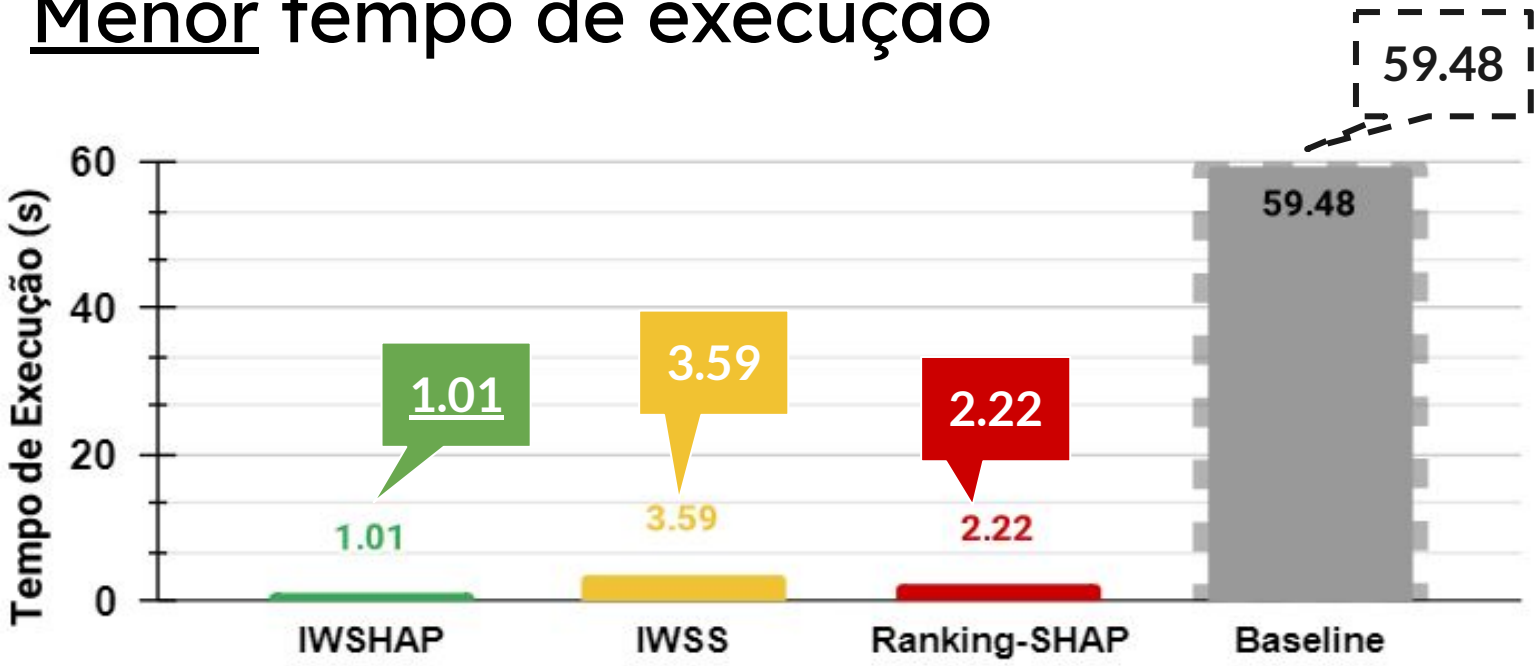
Resultados

- IWSHAP <u>menor quantia de caracteristicas</u>



Resultados

- Menor tempo de execução

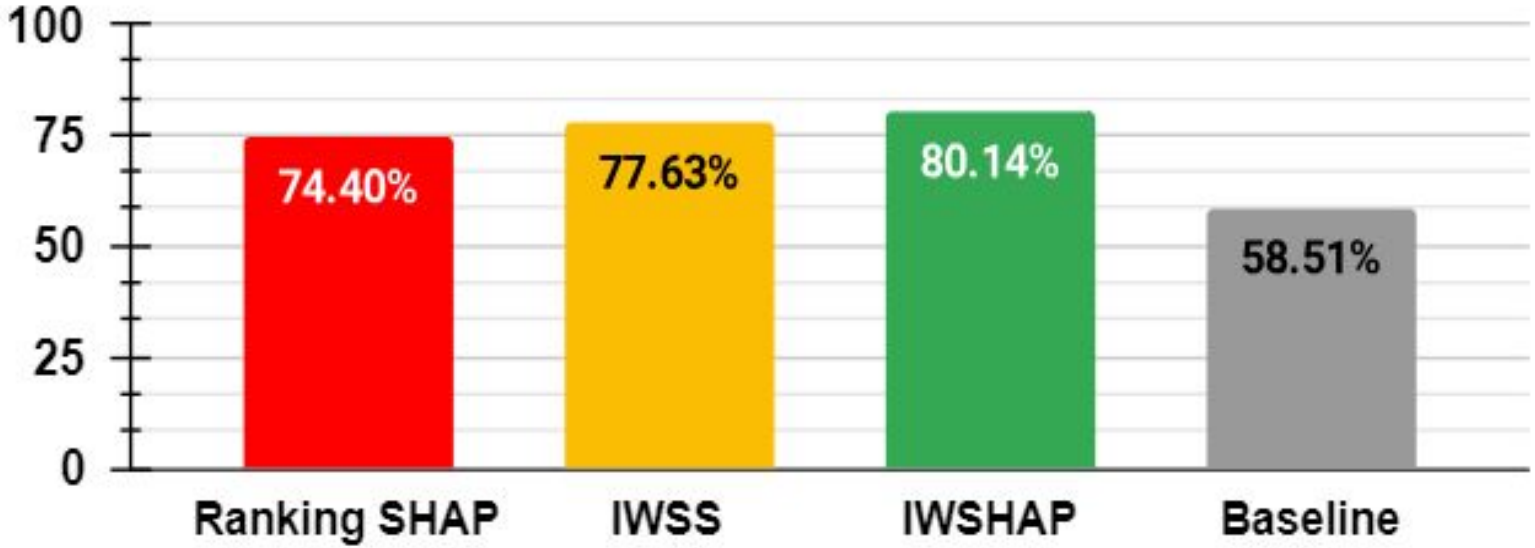


Resultados

- Principais contribuições
  - Aumentou as métricas
  - Reduziu em 99,17% as características
  - Reduziu em 98,3% o tempo de execução

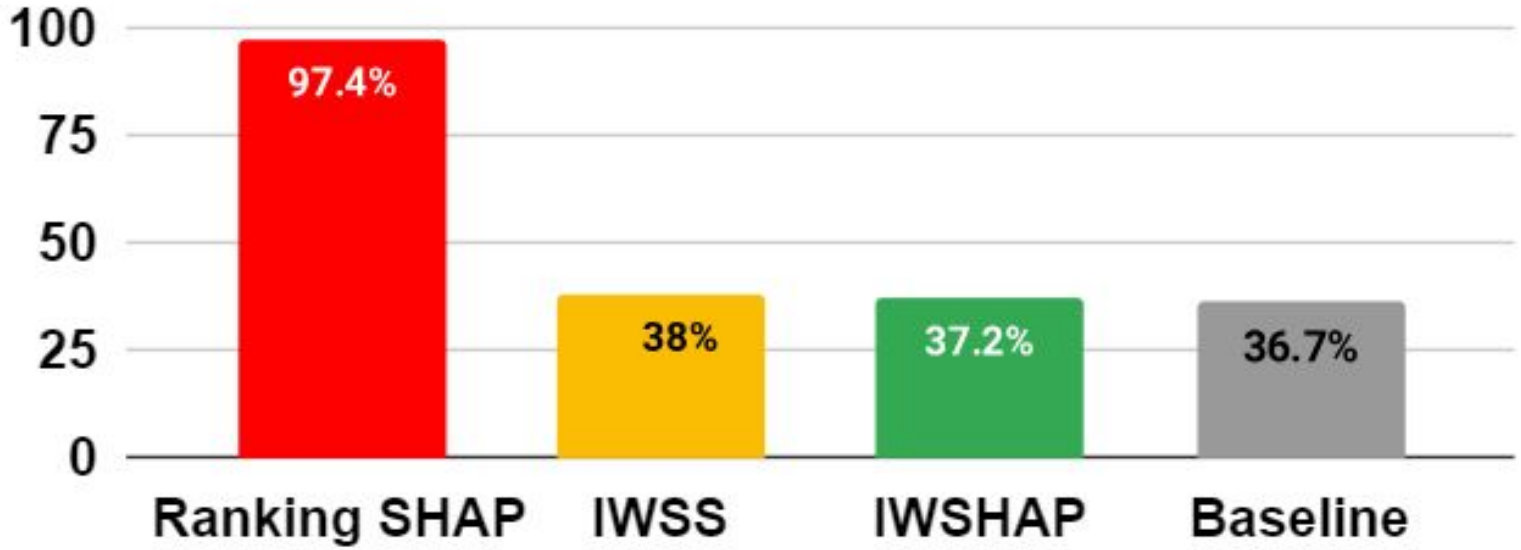
Uso Computacional

- Uso de CPU



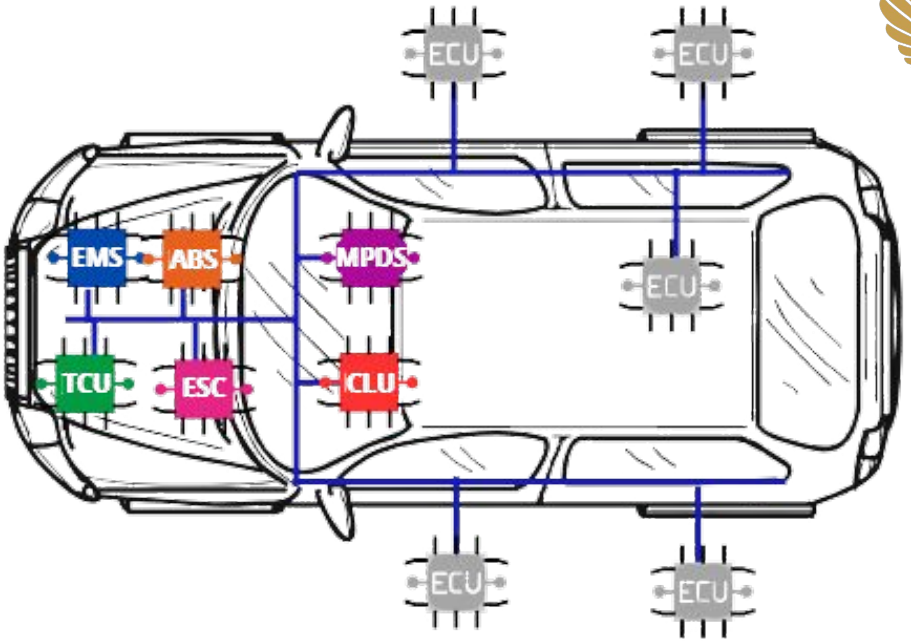
Uso Computacional

- <u>Eficiência</u> no uso de memória



Análise

- Interpretação de Ataques Com o IWSHAP



Unidade de Painel de Instrumentos (CLU)
Controle Eletrônico de Estabilidade (ESC)
Sistema de Freios Antibloqueio (ABS)
Unidade de Controle da Transmissão (TCU)
Sistema de Gerenciamento do Motor (EMS)
Direção Assistida por Motor Elétrico (MDPS)
Outras ECUs (Travas, Suspenção, etc.)

Considerações finais

- O IWSHAP é capaz de:
  - Melhorar as métricas de desempenho!
  - Reduzir a quantidade de características!
  - Reduzir o tempo de execução
    - **99,17% e 98,3%**

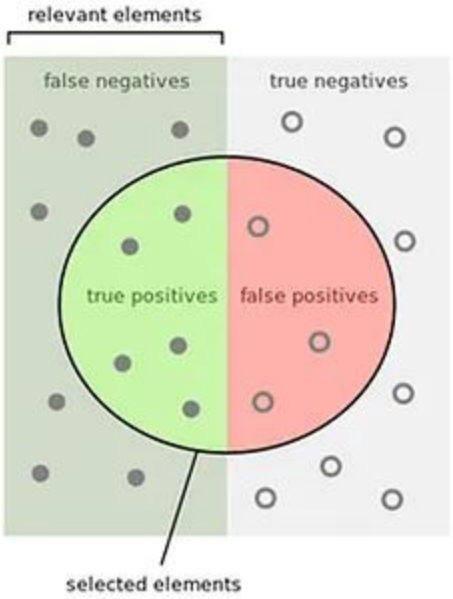
Trabalhos futuros

- Explorar diferentes contextos
- Integrar outras técnicas de XAI
- Explorar diferentes algoritmos de AM e RN
- Implementação em ambientes reais e abordar novos datasets

Obrigado!

Felipe H. Scherer - felipescherer.aluno@unipampa.edu.br

Anexos



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{selected elements}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{relevant elements}}$$

Anexos

- Referências das notícias:
  - cbc.ca/news/science/hackers-kill-engine-of-moving-jeep-on-highway-in-security-demo-1.3162944
  - https://www.theguardian.com/technology/2016/sep/20/tesla-model-s-chinese-hack-remote-control-brakes#:~:text=Three%20months%20since%20the%20first,in%20the%20high%2Dtech%20car.

Referências

- Nazat, S., Li, L., & Abdallah, M. (2024). **XAI-ADS: An explainable artificial intelligence framework for enhancing anomaly detection in autonomous driving systems**. *IEEE Access*, 12, 48583–48607.
- Roshan, K., & Zafar, A. (2021). **Utilizing XAI technique to improve autoencoder based model for computer network anomaly detection with shapley additive explanation (SHAP)**. *International Journal of Computer Networks Communications (IJCNC)*, 13(6), 109–128.
- Setitra, M. A., Fan, M., & Bensalem, Z. E. A. (2023). **An efficient approach to detect distributed denial of service attacks for software defined internet of things combining autoencoder and extreme gradient boosting with feature selection and hyperparameter tuning optimization**. *Transactions on Emerging Telecommunications Technologies*, 34(9), e4827.
- Ullah, S., Khan, M. A., Ahmad, J., Jamal, S. S., Huma, Z., Hassan, M. T., Pitropakis, N., Arshad, & Buchanan, W. J. (2022). **HDL-IDS: A hybrid deep learning architecture for intrusion detection in the Internet of Vehicles**. *Sensors*, 22(4).

Referências

- Asry, C. E. L., Benchaji, I., Douzi, S., & Ouahidi, B. E. L. (2024). **A robust intrusion detection system based on a shallow learning model and feature extraction techniques**. *PLOS ONE*, 19(1), 1–31.
- Aksu, D., & Aydin, M. A. (2022). **MGA-IDS: Optimal feature subset selection for anomaly detection framework on in-vehicle networks-CAN bus based on genetic algorithm and intrusion detection approach**. *Computers & Security*, 118, 102717.
- Bhandari, S., Kukreja, A. K., Lazar, A., Sim, A., & Wu, K. (2020). **Feature selection improves tree-based classification for wireless intrusion detection**. In *Proceedings of the 3rd International Workshop on Systems and Network Telemetry and Analytics, SNTA '20*, 19–26. New York, NY, USA: Association for Computing Machinery.
- Mowla, N. I., Rosell, J., & Vahidi, A. (2022). **Dynamic voting based explainable intrusion detection system for in-vehicle network**. In *2022 24th International Conference on Advanced Communication Technology (ICACT)*, 406–411.