HACKERS DO BEM

GT Malware DataLab

https://malwaredatalab.github.io

Demo Day

29 de novembro de 2024

Equipe

Diego Kreutz

Experiência em Projetos de Malware Android e formação de recursos humanos

Rodrigo Mansilha

Experiência em Projetos de IA generativa e formação de recursos humanos

Hendrio Bragança

Experiência em IA, Ciência de Dados e Soluções para malware Android

Kayuã Paim

Experiência em IA Generativa e Desenvolvimento de Software

Luiz Laviola

Experiência em Sistemas, Engenharia e Desenvolvimento de Software

Angelo Diniz

Experiência em Sistemas, Infraestrutura e Desenvolvimento de Software

Colaboradores

Anna Luiza Gomes
Profissional de Ciência de Dados
com Experiência de Mercado
Mestranda em Eng. de Software

Leonardo Sonco
Profissional de Front End
Graduando em Ciência da
Computação

Lucas Ferreira
Profissional de Cibersegurança
com Experiência de Mercado
Mestrando em Eng. de Software

Rafael Nogueira
Voluntário de Iniciação Tecnológica
Graduando em Ciência da
Computação

Malware Android: o problema persiste

HACKERS DO BEM

Android malware is wiping out bank accounts in Finland

Updated on: May 06, 2024 8:55 AM

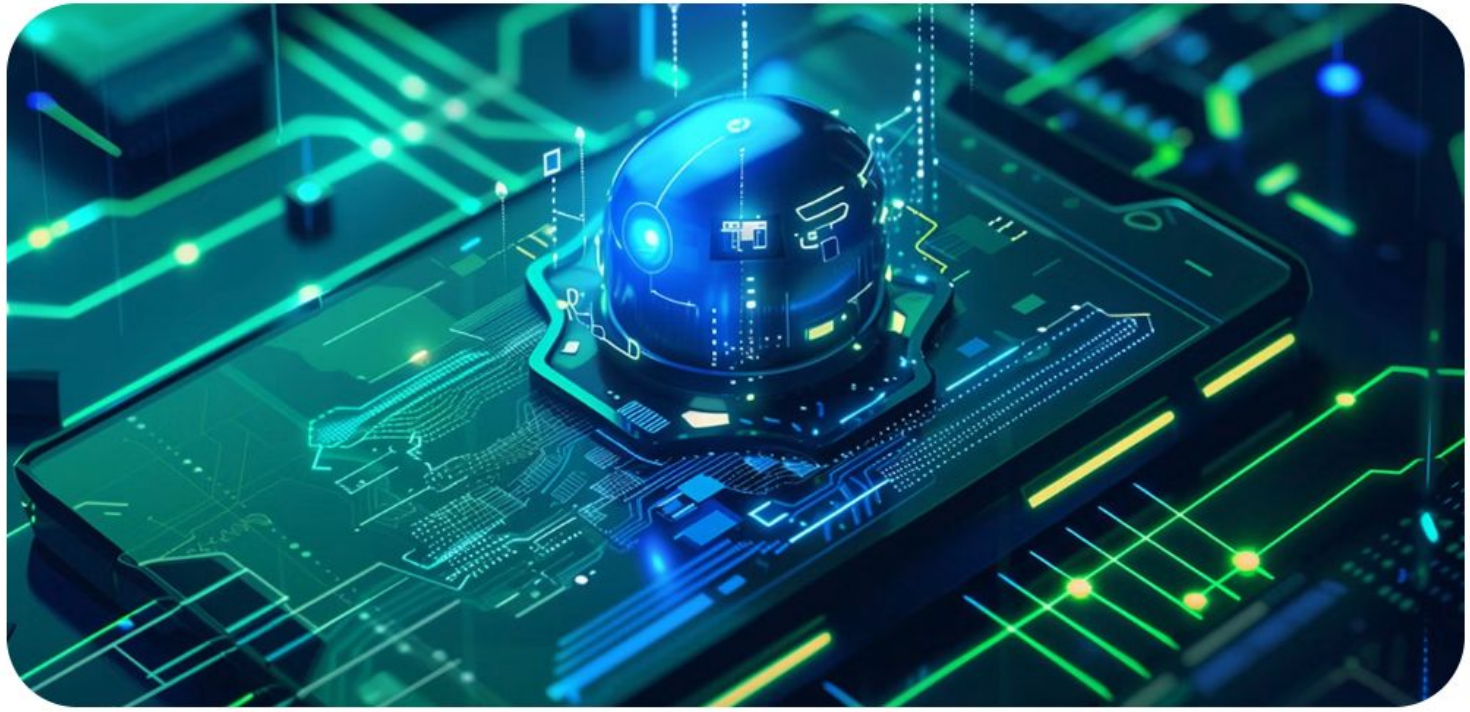
Paulina Okunytė, Journalist

Usuários perdem várias dezenas de milhares de EUROs

Como resolver o problema?

Google Unveils Advanced AI Malware Detection for Android

By Sophia Taylor



Inteligência Artificial (IA) Avançada é necessária para combater malwares

Mas por que projetos de IA falham?

14 Reasons Why 70 to 80% of AI projects fail: Decoding the High Failure Rate of AI Projects

Kenneth Nicholaus · Follow

5 min read · Jul 12, 2024

Dados insuficientes ou de baixa qualidade e falta de cultura orientada a dados. IA é centrada nos dados!
Os datasets existentes são defasados e limitados!

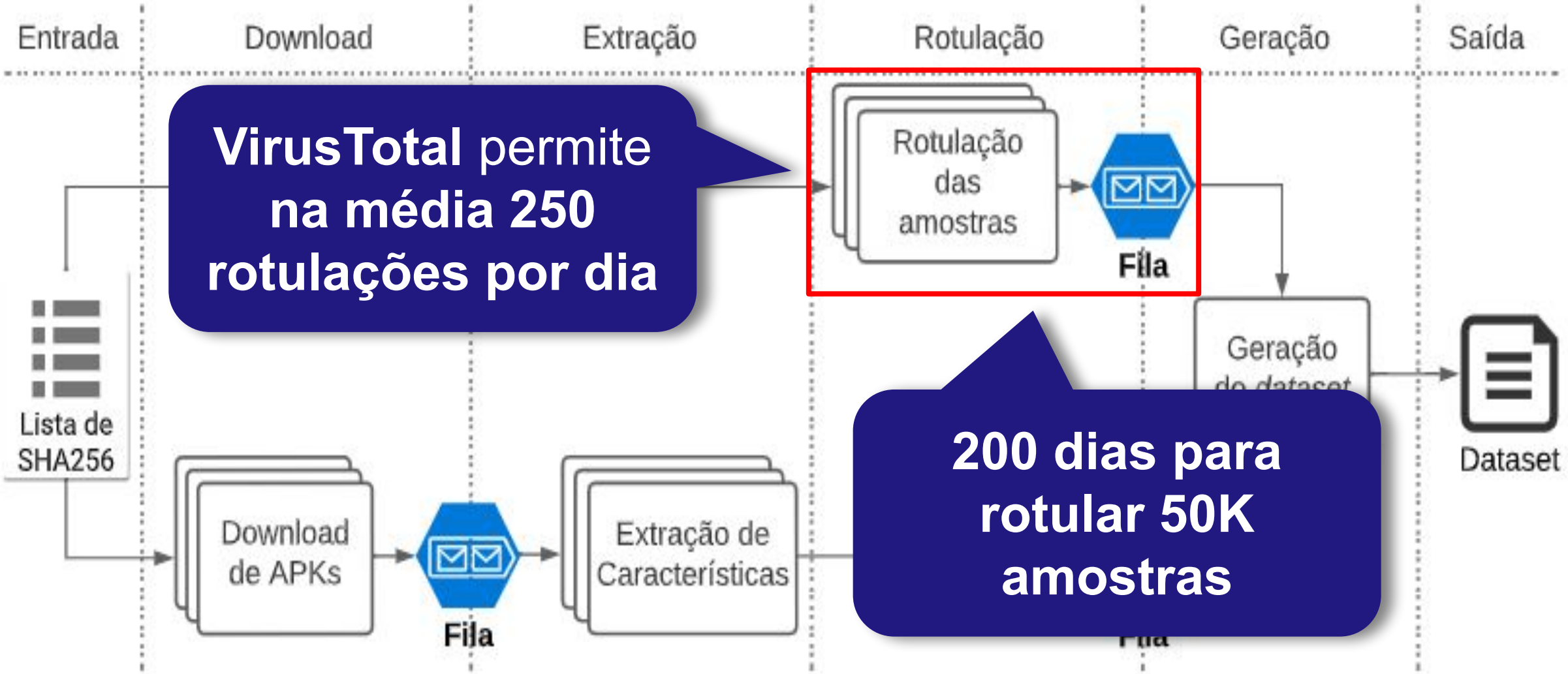
Como obter datasets de qualidade?

HACKERS DO BEM

ADBuilder

AMGenerator e AMExplorer

ADBuiler


Entrada
Download
Extração
Rotulação
Geração
Saída
VirusTotal permite na média 250 rotulações por dia
Rotulação das amostras
Fila
Lista de SHA256
Download de APKs
Fila
Extração de Características
Geração
Dataset
200 dias para rotular 50K amostras

Como resolver o problema?

Forbes

FORBES > INNOVATION > ENTERPRISE TECH

20 Generative AI Tools For Creating Synthetic Data

Bernard Marr Contributor ⓘ

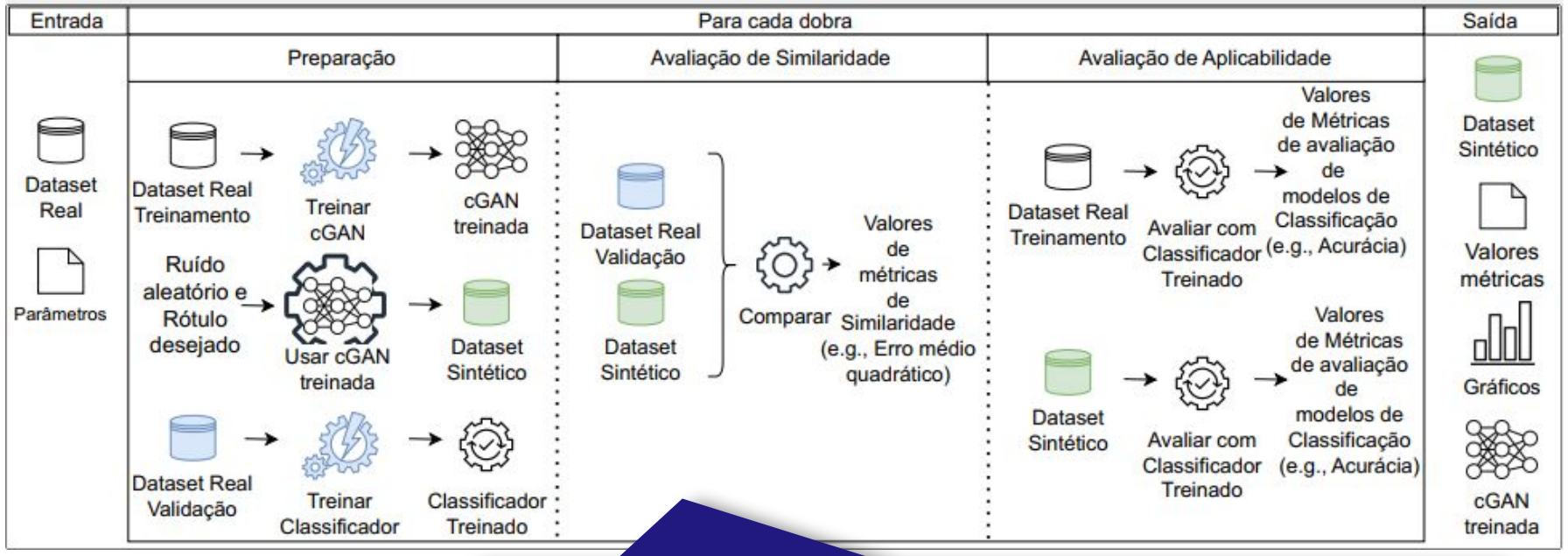
Follow

0

Aug 29, 2024, 02:51am EDT

Uma resposta: Ferramentas de IA Generativa para Geração de Dados Sintéticos

DroidAugmentor (2023)

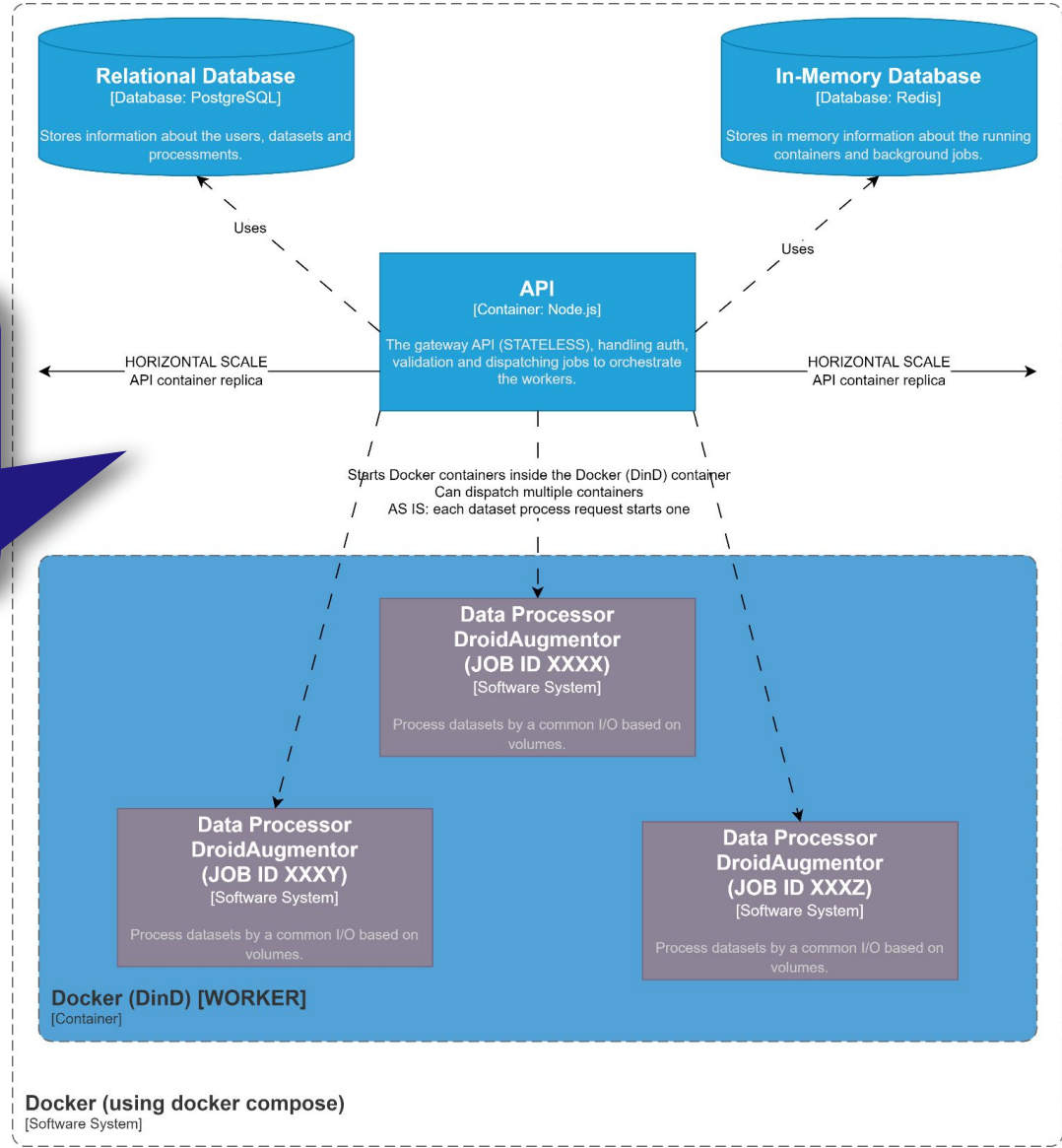


1. **preparação**: treinamento da cGAN
2. **avaliação** (**similaridade**): métricas
3. **avaliação** (**aplicabilidade**): métricas

DroidAugmentor: uma ferramenta de treinamento e avaliação de cGANs para a geração de dados sintéticos
Karina Casola, Kayuã Oleques Paim, Diego Kreutz, Rodrigo Brandão Mansilha
XXIII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais
https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/27273

AutoDroid (2023)

- Código Open Source (GitHub)
- Desenvolvido em Typescript para Node.js
- Utilização através de API REST



AutoDroid: disponibilizando a ferramenta DroidAugmentor como serviço
Luiz Felipe Laviola, Kayuã Oleques Paim, Diego Kreutz, Rodrigo Brandão Mansilha
XX Escola Regional de Redes de Computadores (ERRC 2023)
https://sol.sbc.org.br/index.php/errc/article/view/26020

Hackers do Bem (2023)

GT - Malware DataLab: Serviço de Execução e Avaliação de Redes Neurais Artificiais para Geração de Dados Sintéticos de Malware

GT - Malware DataLab (2023)

Laboratório de Experimentação (produto 1)

Oferecer um serviço para reduzir a curva de aprendizado e facilitar a investigação de técnicas avançadas de geração de dados sintéticos para expandir *datasets* úteis na classificação de aplicativos Android (benigno ou maligno).

O1. Ambiente de Execução (extensão inovadora)

O2. Ambiente de Avaliação (pesquisa)

O3. Material de Aprendizado e Treinamento (ensino)

GT - Malware DataLab (2024)

Laboratório de Ensino-Aprendizagem (produto 2)

Capacitar os Hackers do Bem através de uma plataforma de Ensino-Aprendizagem para formação de recursos humanos qualificados em IA Generativa para gerar dados sintéticos.

- Módulos com material teórico e prático
- Avaliação diagnóstica e avaliações somativas
- Introdução a laboratórios adicionais de aprendizagem

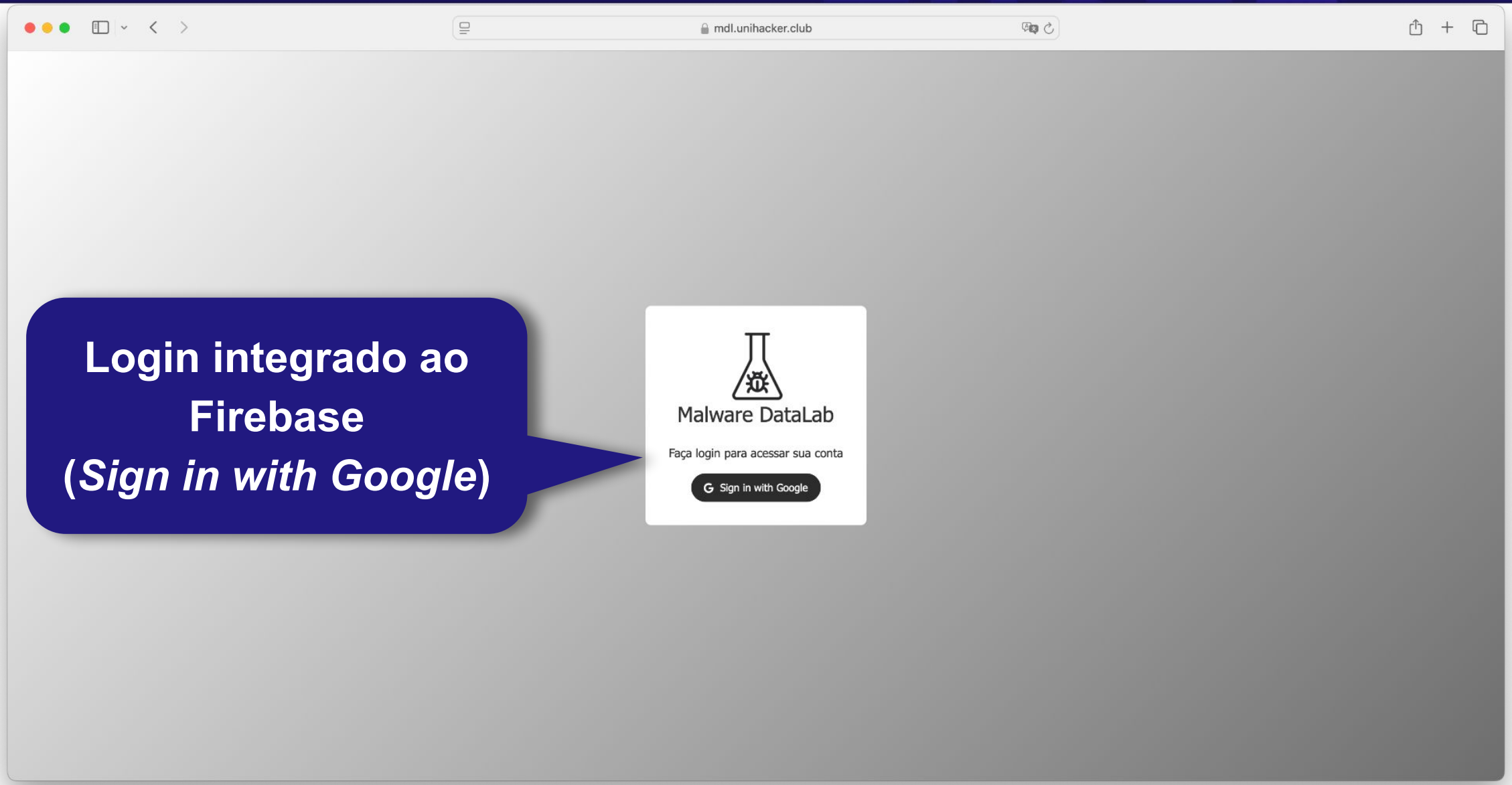
Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. Laboratório de Experimentação
3. MalSynGen (evolução da DroidAugmentor)
4. Cloud AutoDroid (evolução da AutoDroid)
5. GUI4MalSynGen (versão minimalista)

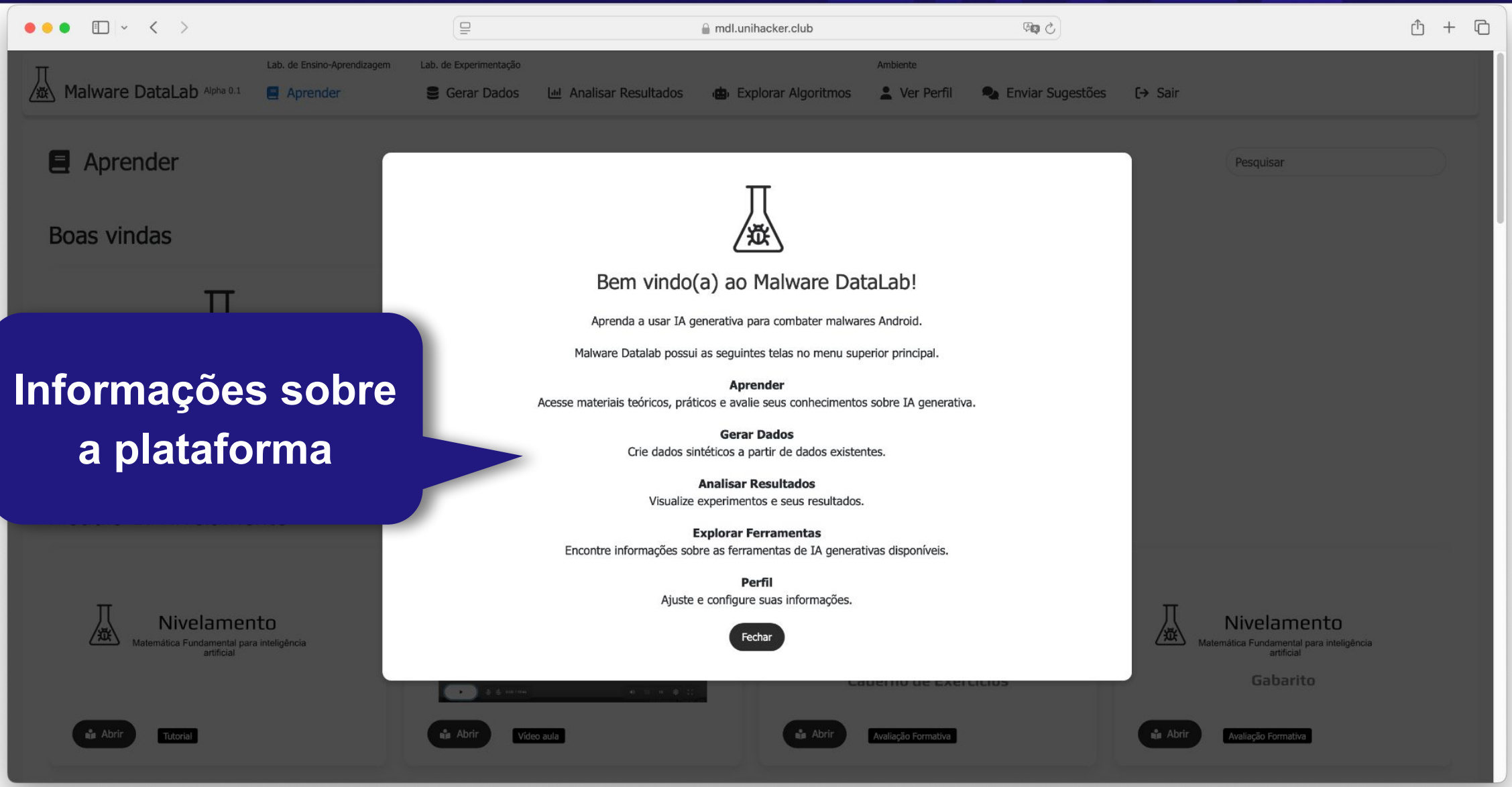
Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. Laboratório de Experimentação
3. MalSynGen (evolução da DroidAugmentor)
4. Cloud AutoDroid (evolução da AutoDroid)
5. GUI4MalSynGen (versão minimalista)

Malware DataLab: Login

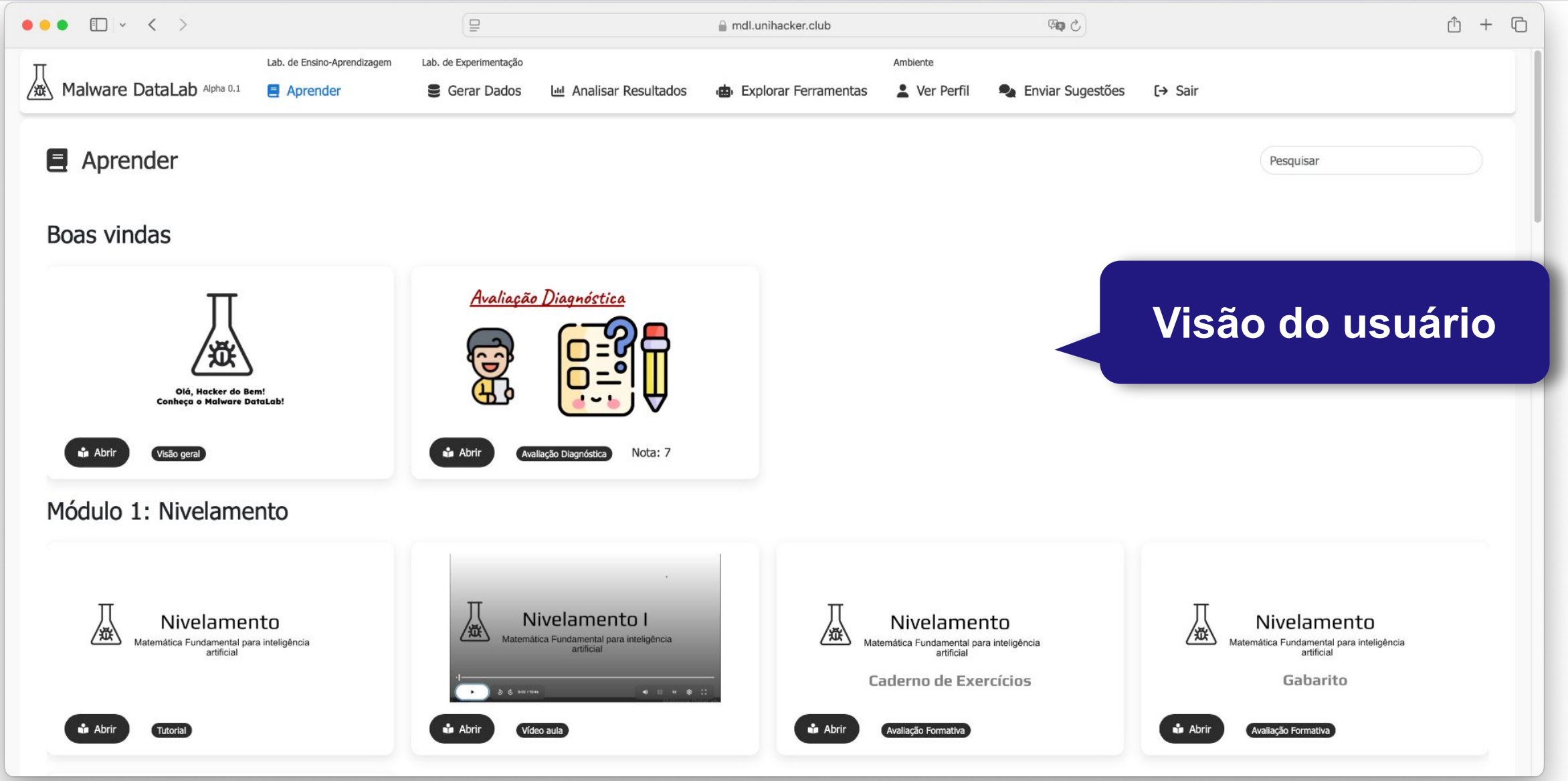


Malware DataLab: Boas vindas



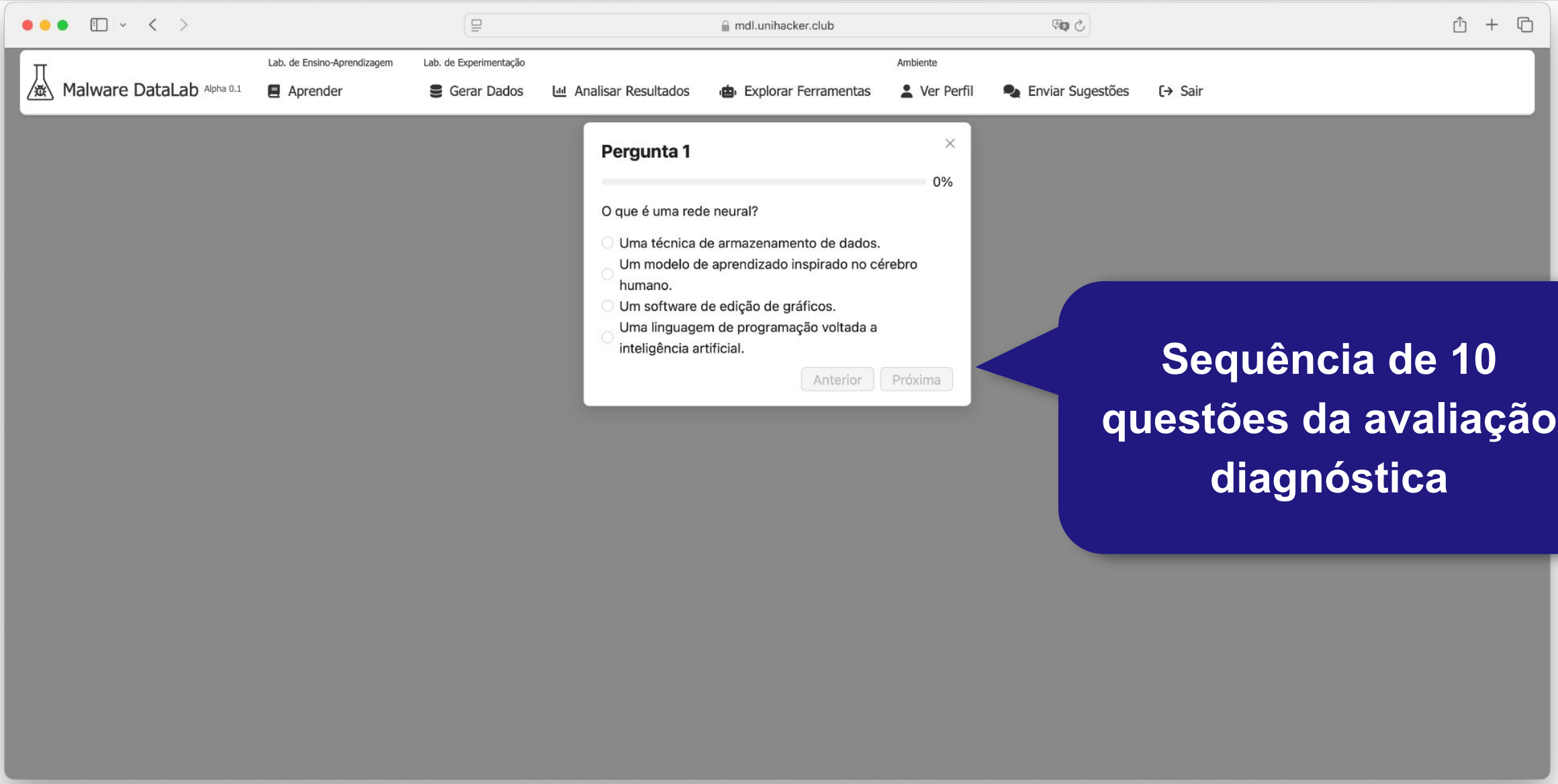
Informações sobre a plataforma

Lab. de Ensino-Aprendizagem



Visão do usuário

Avaliação diagnóstica



Sequência de 10 questões da avaliação diagnóstica

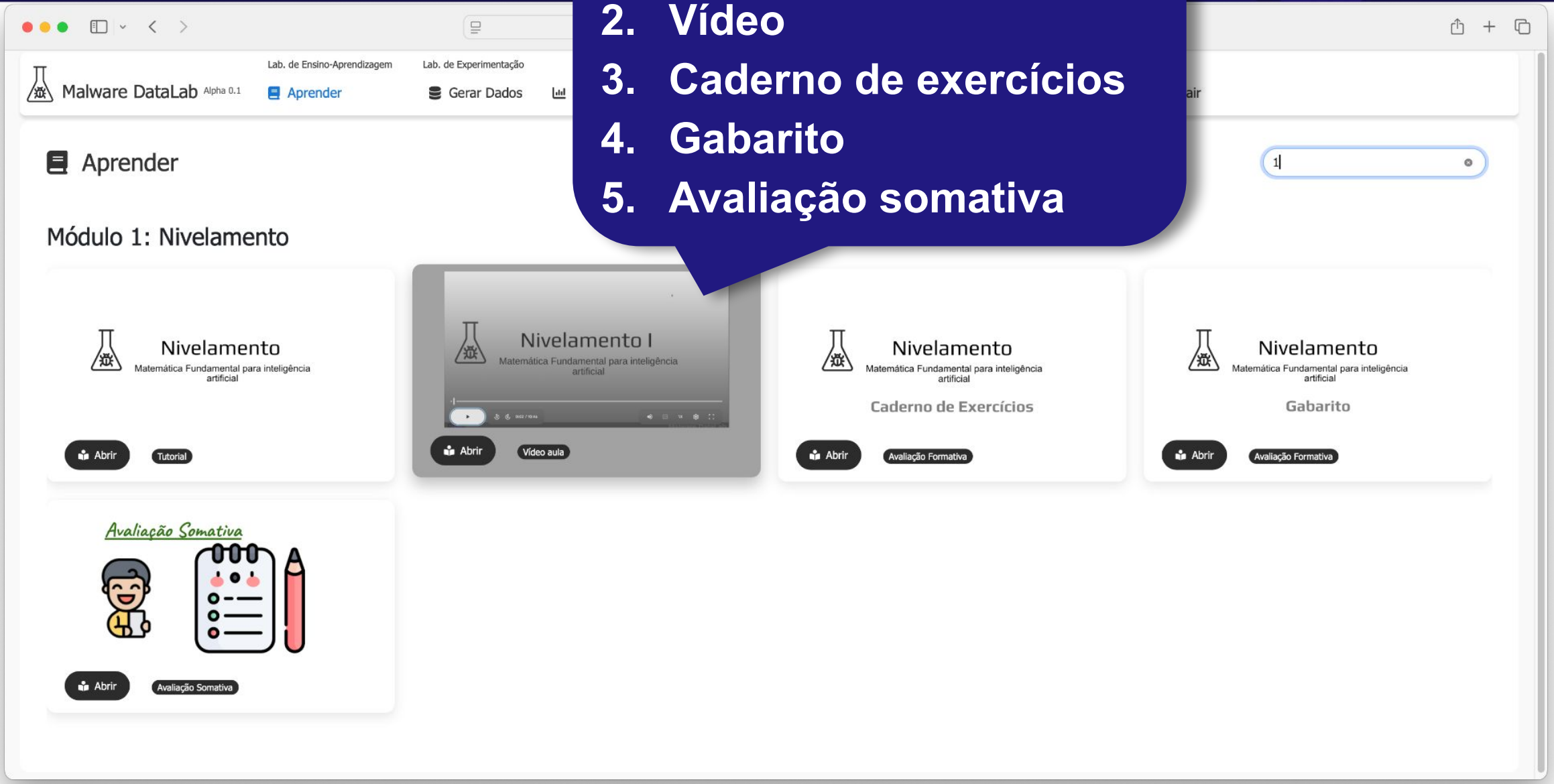
Diagnóstico

Resultado e recomendação da avaliação diagnóstica

Módulos

Módulo 1: Nivelamento

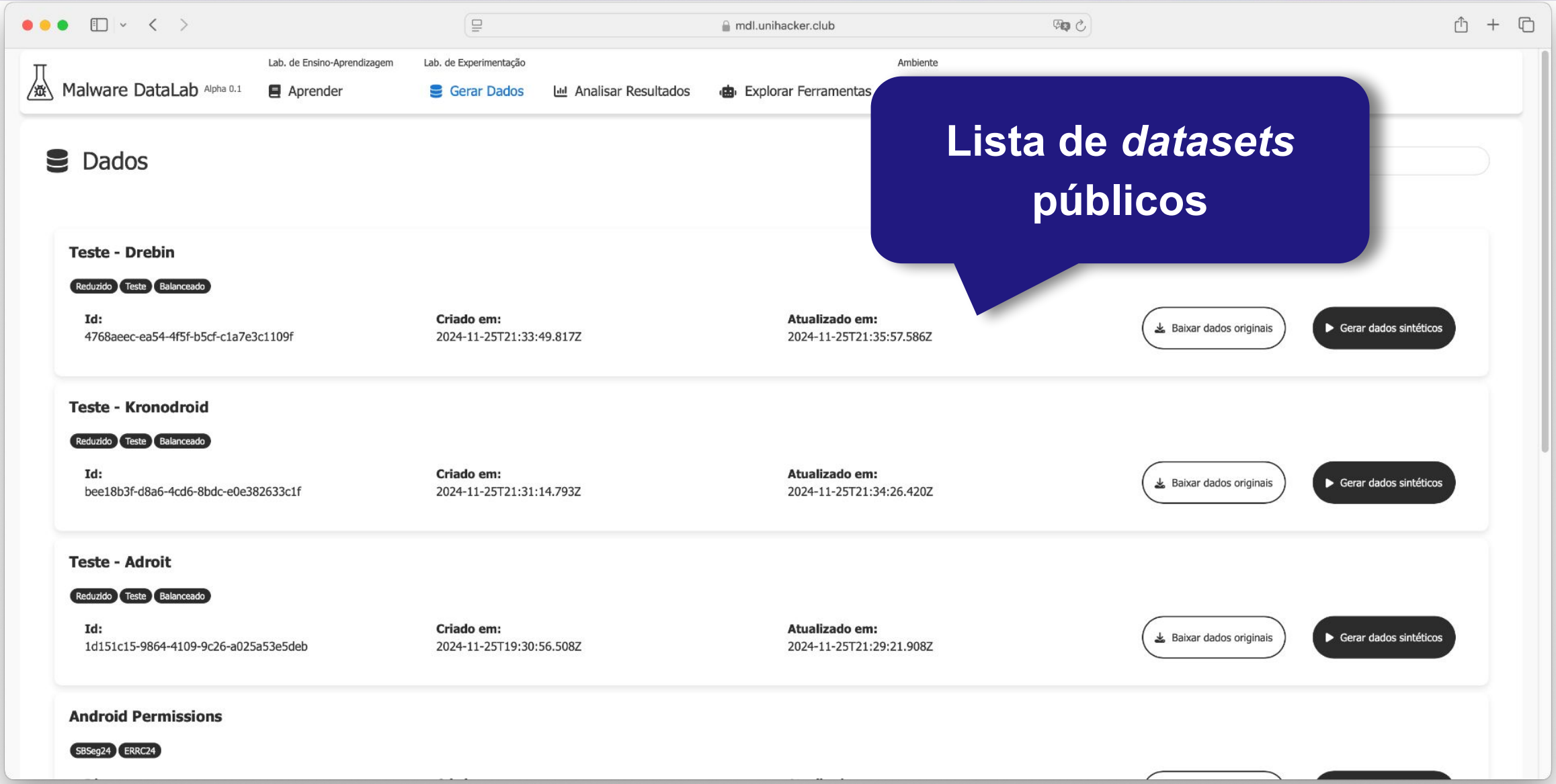
1. Slides
2. Vídeo
3. Caderno de exercícios
4. Gabarito
5. Avaliação somativa



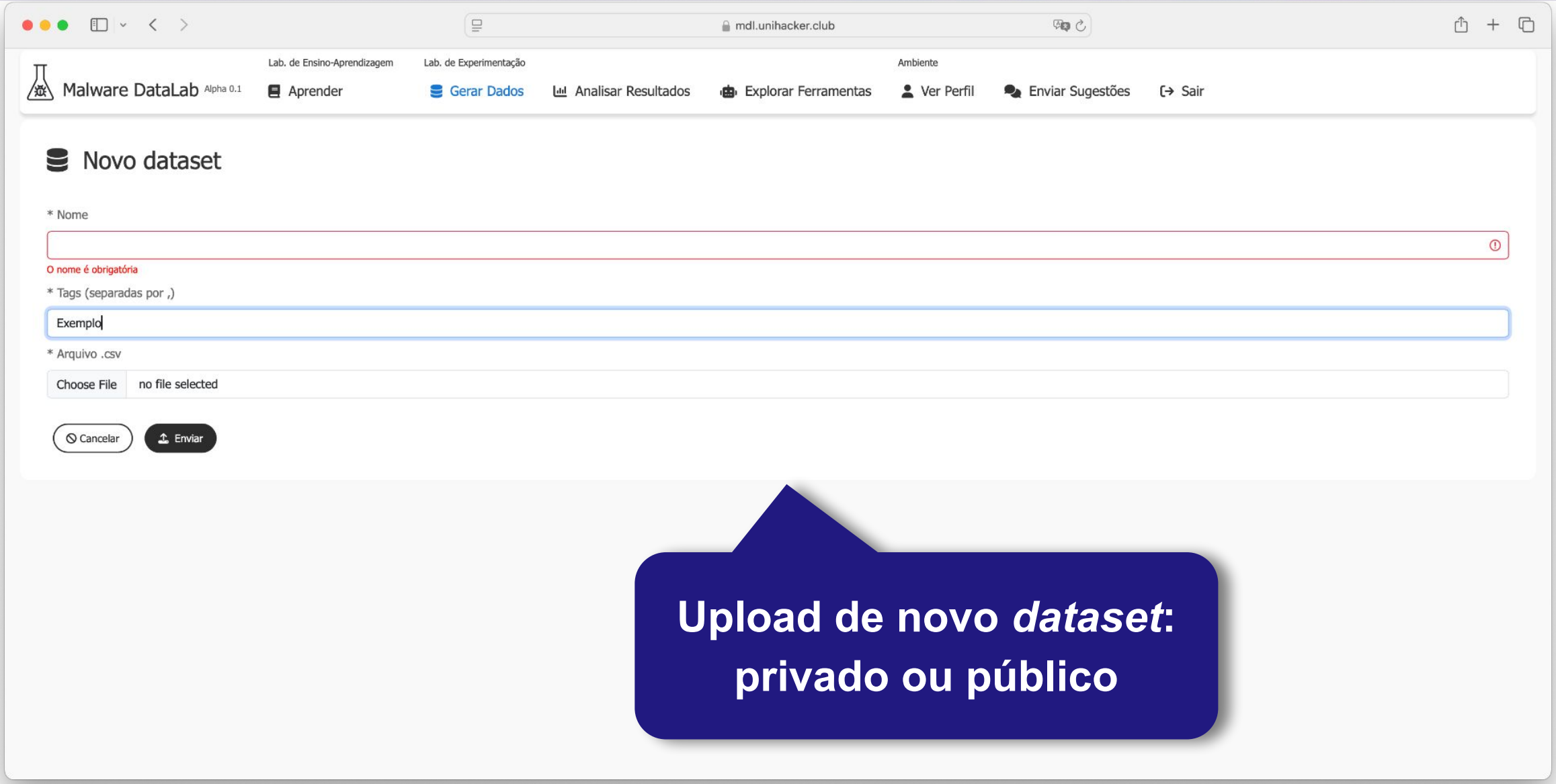
Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. **Laboratório de Experimentação**
3. MalSynGen (evolução da DroidAugmentor)
4. Cloud AutoDroid (evolução da AutoDroid)
5. GUI4MalSynGen (versão minimalista)

Lab. de Experimentação: Dados

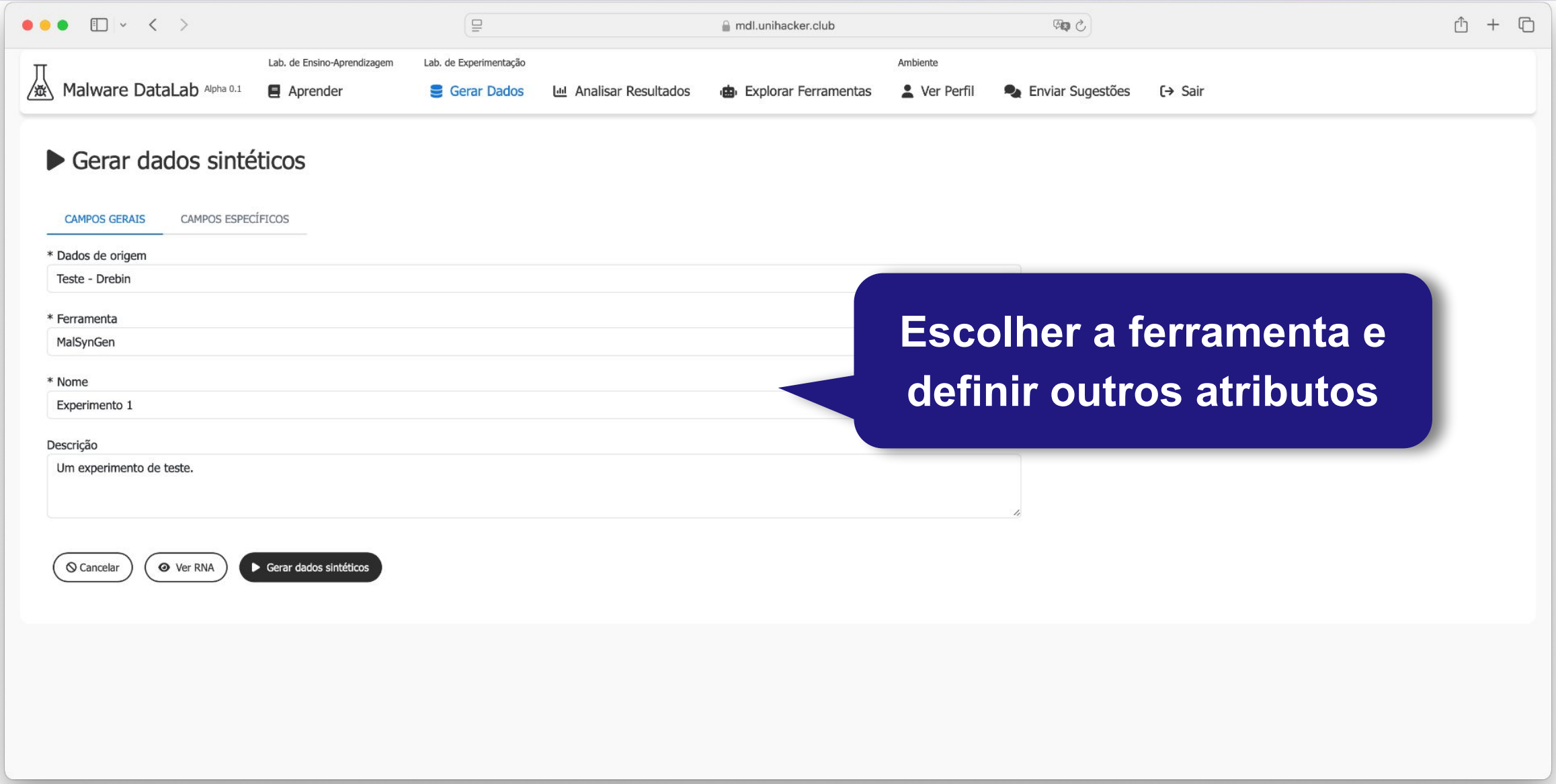


Subir dados

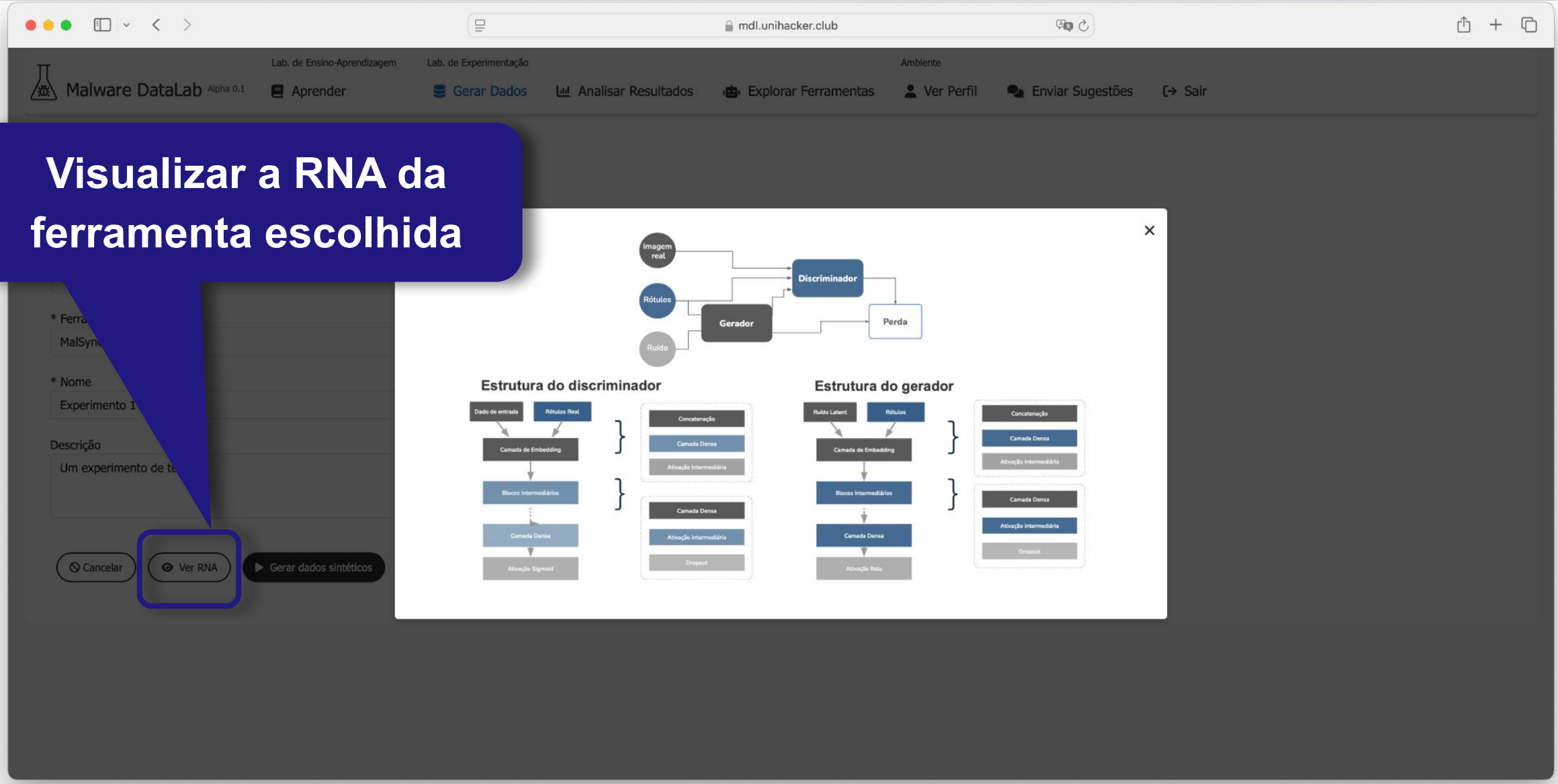


Upload de novo *dataset*: privado ou público

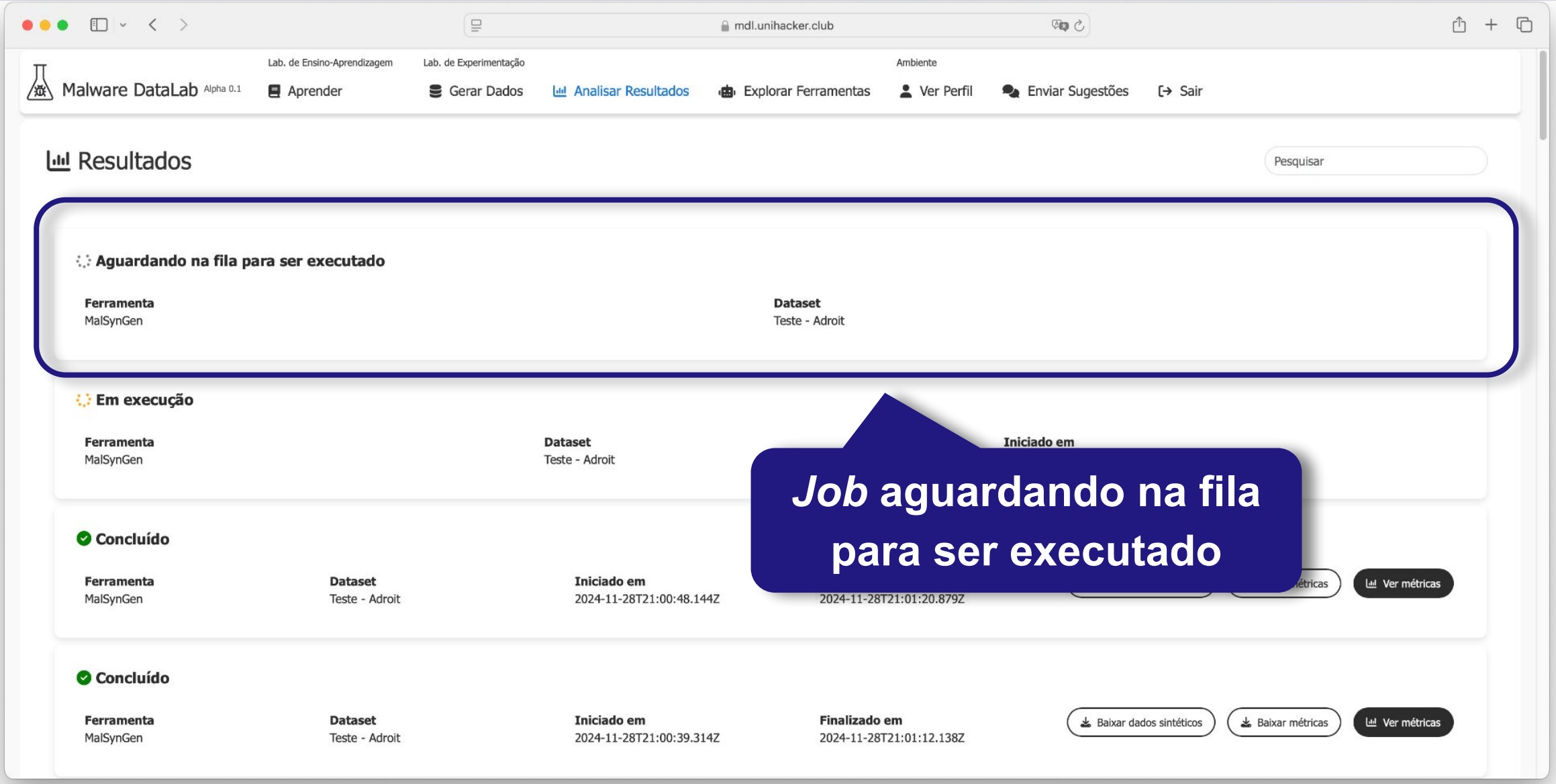
Gerar dados


mdl.unihacker.club
Malware DataLab Alpha 0.1
Lab. de Ensino-Aprendizagem
Aprender
Lab. de Experimentação
Gerar Dados
Analisar Resultados
Explorar Ferramentas
Ambiente
Ver Perfil
Enviar Sugestões
Sair
Gerar dados sintéticos
CAMPOS GERAIS
CAMPOS ESPECÍFICOS
* Dados de origem
Teste - Drebin
* Ferramenta
MalSynGen
* Nome
Experimento 1
Descrição
Um experimento de teste.
Cancelar
Ver RNA
Gerar dados sintéticos
Escolher a ferramenta e definir outros atributos

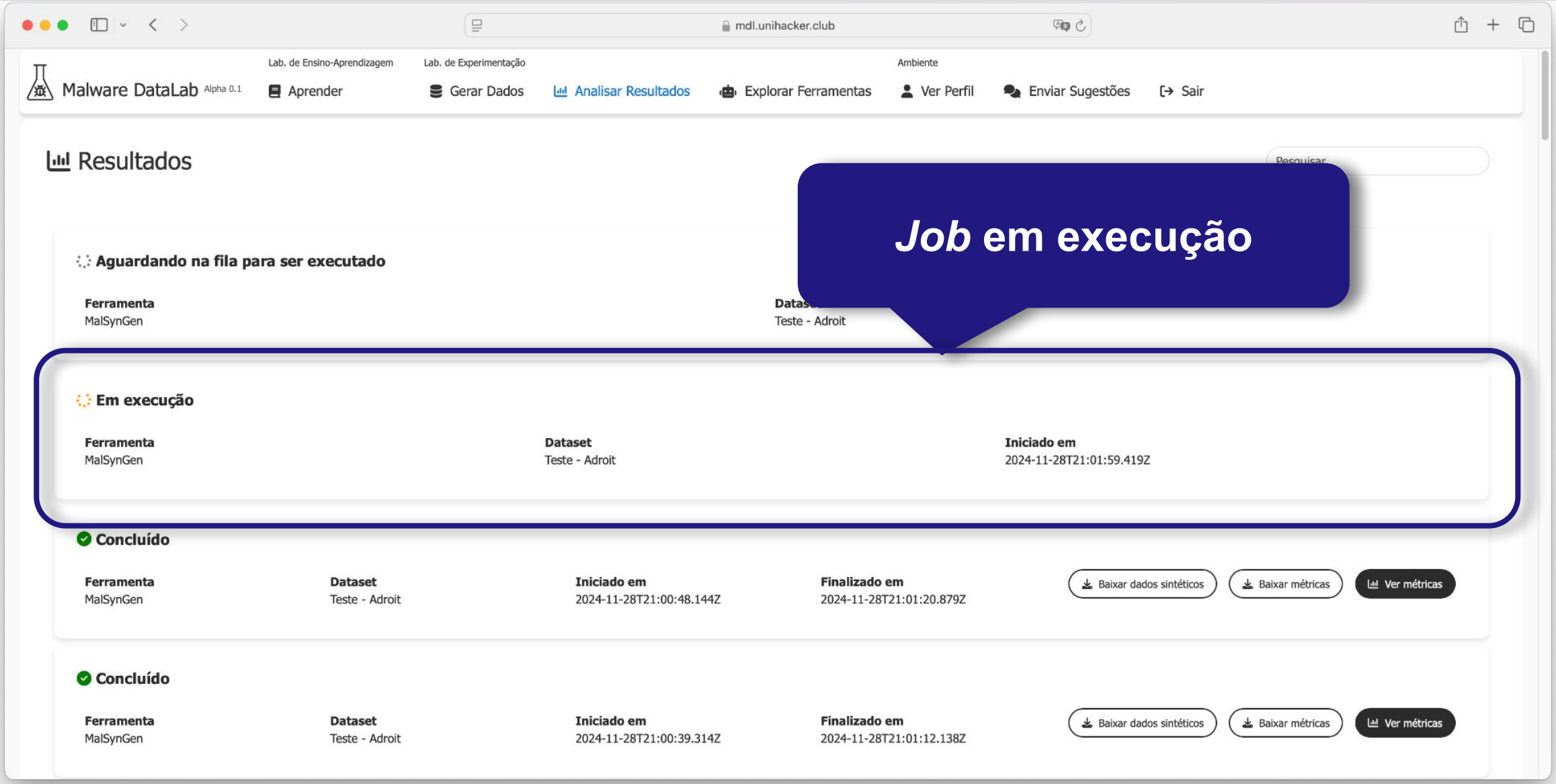
Ver RNA


mdl.unihacker.club
Malware DataLab Alpha 0.1
Lab. de Ensino-Aprendizagem
Aprender
Lab. de Experimentação
Gerar Dados
Analisar Resultados
Explorar Ferramentas
Ambiente
Ver Perfil
Enviar Sugestões
Sair
Visualizar a RNA da ferramenta escolhida
* Nome
Experimento 1
Descrição
Cancelar
Ver RNA
Gerar dados sintéticos
Imagem real
Rótulos
Ruído
Gerador
Discriminador
Perda
Estrutura do discriminador
Estrutura do gerador
Dado de entrada
Rótulos Real
Camada de Embedding
Blocos Intermediários
Camada Densa
Ativação Sigmoid
Concatenação
Ativação Intermediária
Dropout
Ruído Latent
Rótulos
Ativação Relu

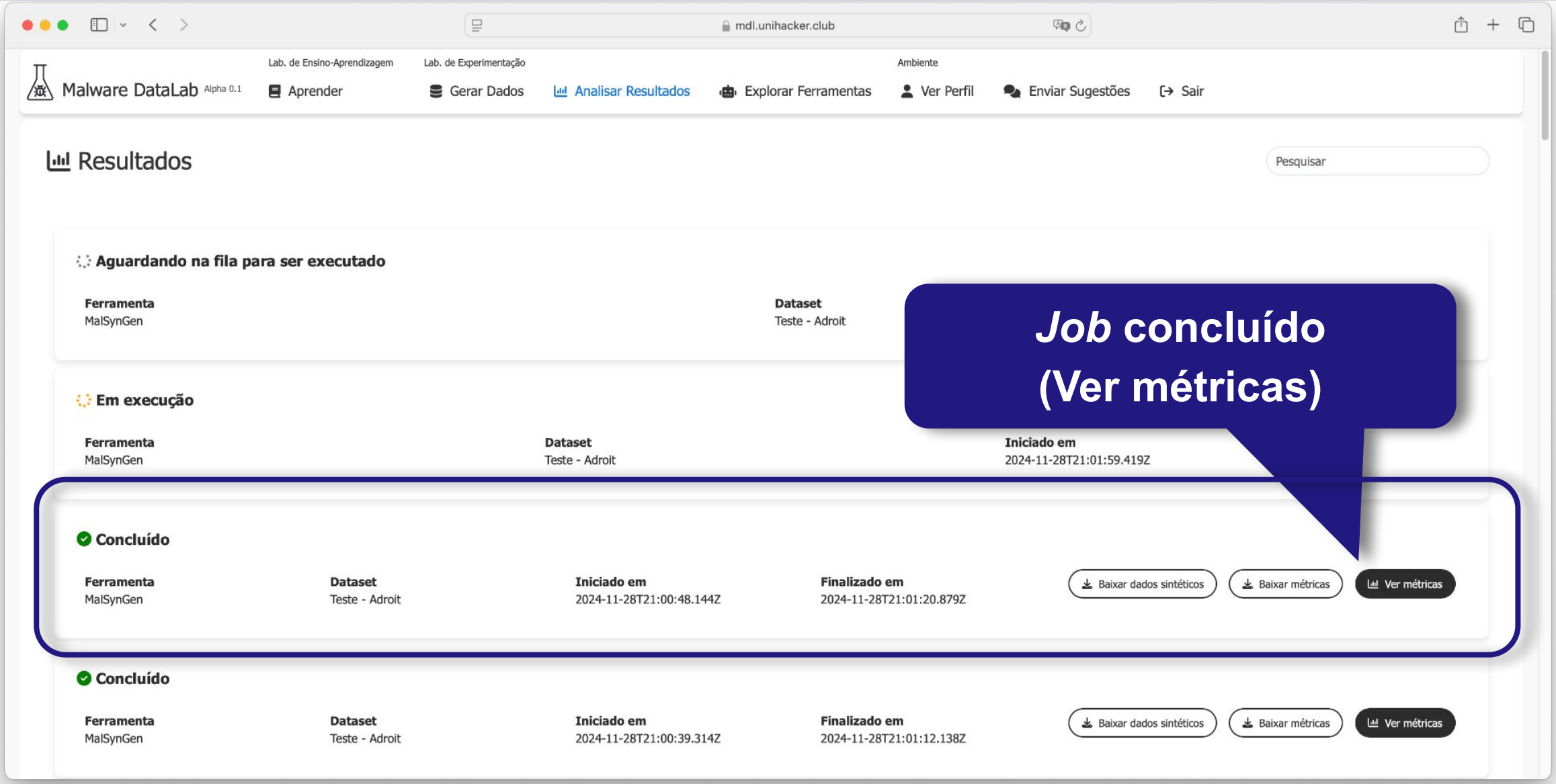
Execução


mdl.unihacker.club
Malware DataLab Alpha 0.1
Lab. de Ensino-Aprendizagem
Aprender
Lab. de Experimentação
Gerar Dados
Analisar Resultados
Explorar Ferramentas
Ambiente
Ver Perfil
Enviar Sugestões
Sair
Resultados
Pesquisar
Aguardando na fila para ser executado
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Em execução
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:48.144Z
2024-11-28T21:01:20.879Z
Ver métricas
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:39.314Z
Finalizado em
2024-11-28T21:01:12.138Z
Baixar dados sintéticos
Baixar métricas
Ver métricas
Job aguardando na fila para ser executado

Execução


mdl.unihacker.club
Malware DataLab Alpha 0.1
Lab. de Ensino-Aprendizagem
Aprender
Lab. de Experimentação
Gerar Dados
Analisar Resultados
Explorar Ferramentas
Ambiente
Ver Perfil
Enviar Sugestões
Sair
Resultados
Pesquisar
Job em execução
Aguardando na fila para ser executado
Ferramenta
MalSynGen
Teste - Adroit
Em execução
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:01:59.419Z
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:48.144Z
Finalizado em
2024-11-28T21:01:20.879Z
Baixar dados sintéticos
Baixar métricas
Ver métricas
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:39.314Z
Finalizado em
2024-11-28T21:01:12.138Z
Baixar dados sintéticos
Baixar métricas
Ver métricas

Execução


mdl.unihacker.club
Malware DataLab Alpha 0.1
Lab. de Ensino-Aprendizagem
Aprender
Lab. de Experimentação
Gerar Dados
Analisar Resultados
Explorar Ferramentas
Ambiente
Ver Perfil
Enviar Sugestões
Sair
Resultados
Pesquisar
Aguardando na fila para ser executado
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Em execução
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:01:59.419Z
Job concluído
(Ver métricas)
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:48.144Z
Finalizado em
2024-11-28T21:01:20.879Z
Baixar dados sintéticos
Baixar métricas
Ver métricas
Concluído
Ferramenta
MalSynGen
Dataset
Teste - Adroit
Iniciado em
2024-11-28T21:00:39.314Z
Finalizado em
2024-11-28T21:01:12.138Z
Baixar dados sintéticos
Baixar métricas
Ver métricas

Analisar Resultados

HACKERS DO BEM

mdl.unihacker.club

Malware DataLab Alpha 0.1 | Lab. de Ensino-Aprendizagem: Aprender | Lab. de Experimentação: Gerar Dados | Sair

Curvas de Treinamento

Teste - Adroit, sintetizado por MalSynGen

PARÂMETROS | CURVAS DE TREINAMENTO | MÉTRICAS DE SIMILARIDADE | MÉTRICAS DE APLICABILIDADE

Interação entre Gerador

As figuras ilustram a interação entre o gerador e o discriminador durante o

Gerador: tem o objetivo de criar amostras

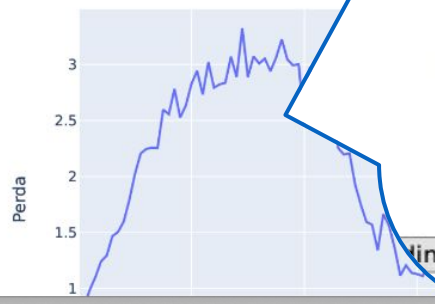
Discriminador: busca melhorar continuamente su

Competição e Convergência: essa competição entre gerador e discriminador resulta em

A não convergência das redes GAN pode ser identificada monitorando as curvas de perda. Estas devem a

Para cada fold (dobra) de trei

Perda do Gerador e Discriminador



ling [MathJax]/extensions/MathMenu.js

Analisar Resultados

mdl.unihacker.club

Malware DataLab Alpha 0.1 | Lab. de Ensino-Aprendizagem | Aprender | Lab. de Experimentação | Gerar Dados | Sair

Métricas de Similaridade

Teste - Adroit, sintetizado por MalSynGen

PARÂMETROS | CURVAS DE TREINAMENTO | MÉTRICAS DE SIMILARIDADE | MÉTRICAS DE APLICABILIDADE | ...S DE CONFUSÃO

As métricas de similaridade avaliam se os dados gerados diferem dos dados originais,

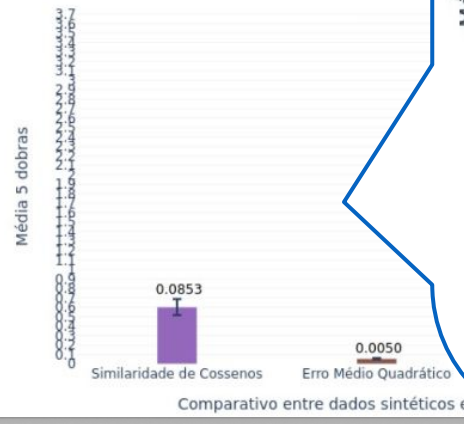
Similaridade de Cossenos: um valor próximo de 1 (ou se aproximando, mas não exatame

Erro Médio Quadrático: um valor próximo de zero indica m

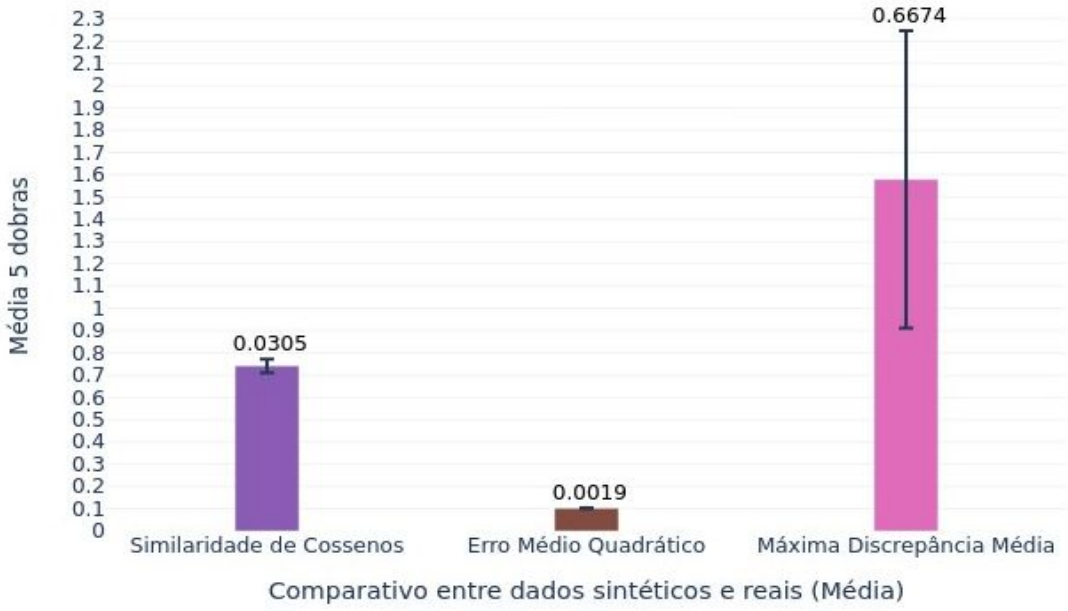
Máxima Discrepância Média: um valor próximo de zero indica

A análise é feita para amostras m

Amostras Ma



Comparativo entre dados sintéticos e re



Amostras Benignas

Analisar Resultados

HACKERS DO BEM

mdl.unihacker.club

Malware DataLab Alpha 0.1 | Lab. de Ensino-Aprendizagem: Aprender | Lab. de Experimentação: Gerar Dados | Sair

Métricas de Aplicabilidade

Teste - Adroit, sintetizado por MalSynGen

PARÂMETROS | CURVAS DE TREINAMENTO | MÉTRICAS DE SIMILARIDADE | MÉTRICAS DE APLICABILIDADE | MATRIZES DE CONFUSÃO

As Métricas de Aplicabilidade avaliam se os classificadores conseguem classificar os dados sintéticos de forma simi

Abordagens Utilizad

TR-AS (Train on Real, Test on Synthetic - TRAs): o modelo é Treinado com

TS-AR (Train on Synthetic, Test on Real - TSAr): o modelo é Treinado com

Para cada abordagem é apresentado um gráfico para ca

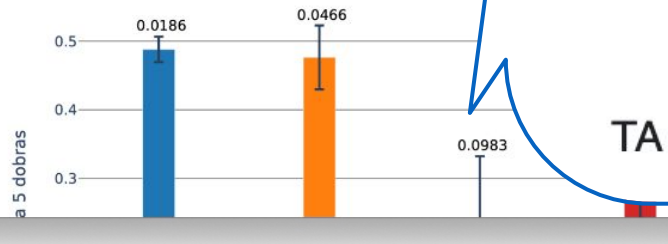
Interpretação dos Resu

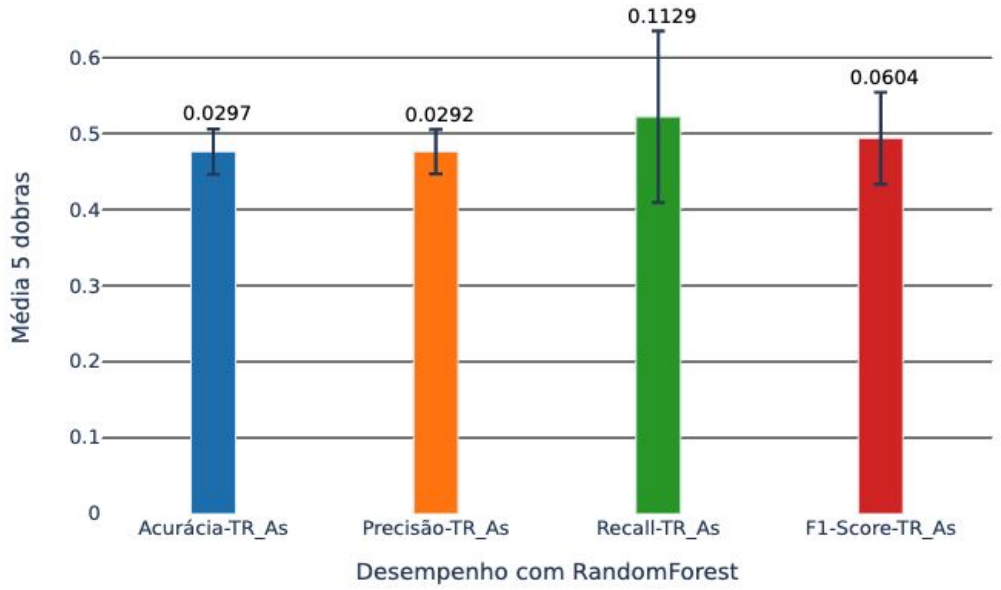
Para métricas de classificação binária:

Quanto mais próximo de 1, melhor é o desempenho do classificador ao utilizar os dados sintéticos. A métrica **recall** é considera

RANDOMFOREST | SUPPORTVECTORMACHIN

TR As: Treinado com Real e Aval





TA Ar: Treinado com Sintético e Avaliado com Real

Analisar Resultados

HACKERS DO BEM

mdl.unihacker.club

Malware DataLab Alpha 0.1 | Lab. de Ensino-Aprendizagem: Aprender | Lab. de Experimentação: Gerar Dados | Sair

Matrizes de Confusão

Teste - Adroit, sintetizado por MalSynGen

PARÂMETROS | CURVAS DE TREINAMENTO | MÉTRICAS DE SIMILARIDADE | MÉTRICAS DE APLICABILIDADE | DE CONFUSÃO

Em geral, as matrizes de confusão são usadas para avaliar os algoritmos de classificação considerando algum dado de entrada. Nós usamos para v
com dados sintéticos se assemelham aos resultados obtidos com os dados reais, m

As matrizes de confusão permitem identificar os acertos e erros do

Falsos Positivos (FP) Falsos Negativos (FN) Verdadeiros Positivos (V

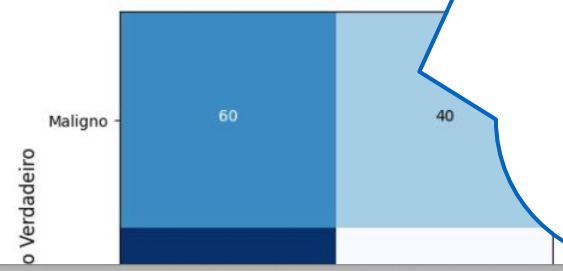
Em cenários específicos, como a classificação entre "benigno" e "maligno", a performance do modelo é avaliad

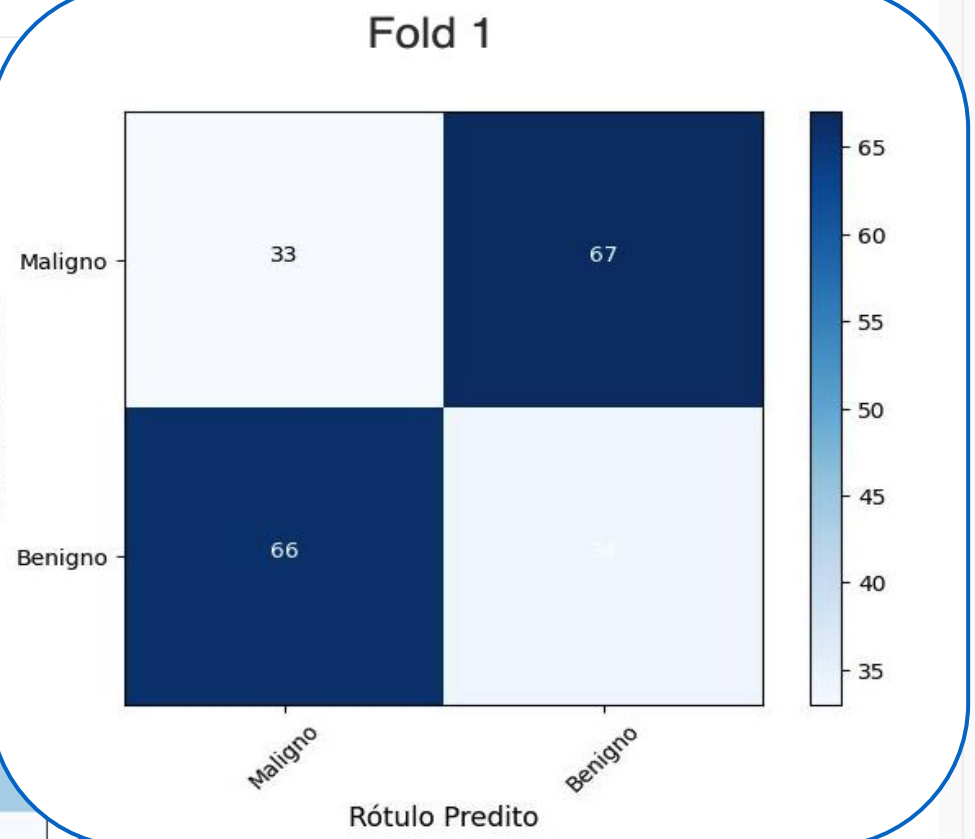
Maligno-Maligno (Verdadeiros Positivos) (quadrante superior esquerdo) **Benigno-Benigno**
Quanto maiores forem esses valores (visualmente, tons mais escuros na diagonal principal), e consequentemente menor os outros valores (visualmer
realizar classificações corretas e confiáveis

RANDOMFOREST | SUPPORTVECTORMACHINE

TR As: Treinado com Real e Avaliado

Fold 1



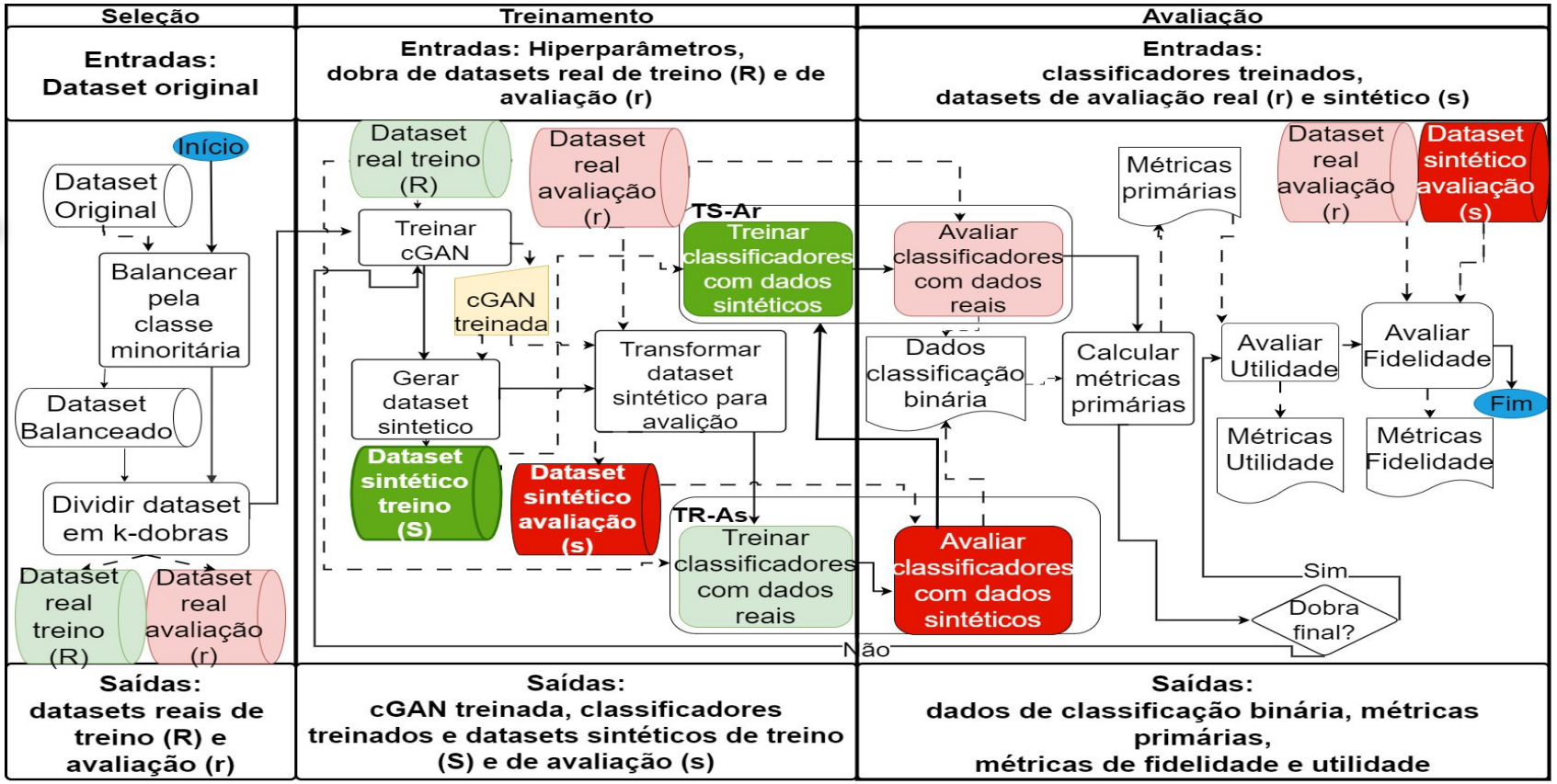


Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. Laboratório de Experimentação
3. **MalSynGen (evolução da DroidAugmentor)**
4. Cloud AutoDroid (evolução da AutoDroid)
5. GUI4MalSynGen (versão minimalista)

MalSynGen

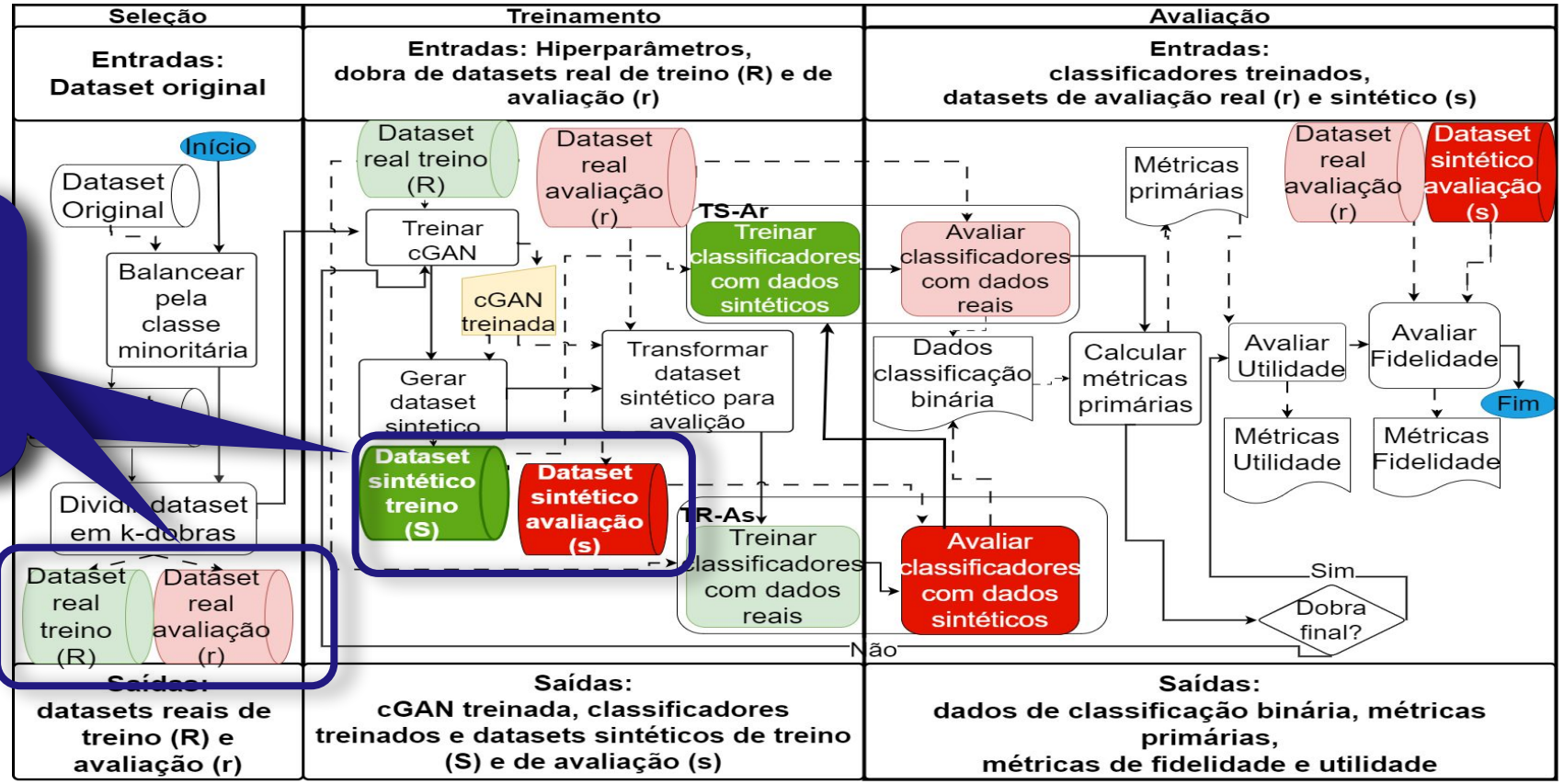
Nova arquitetura e nova metodologia de avaliação



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
VIII Salão de Ferramentas do SBSeg 2024
https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/30126

MalSynGen

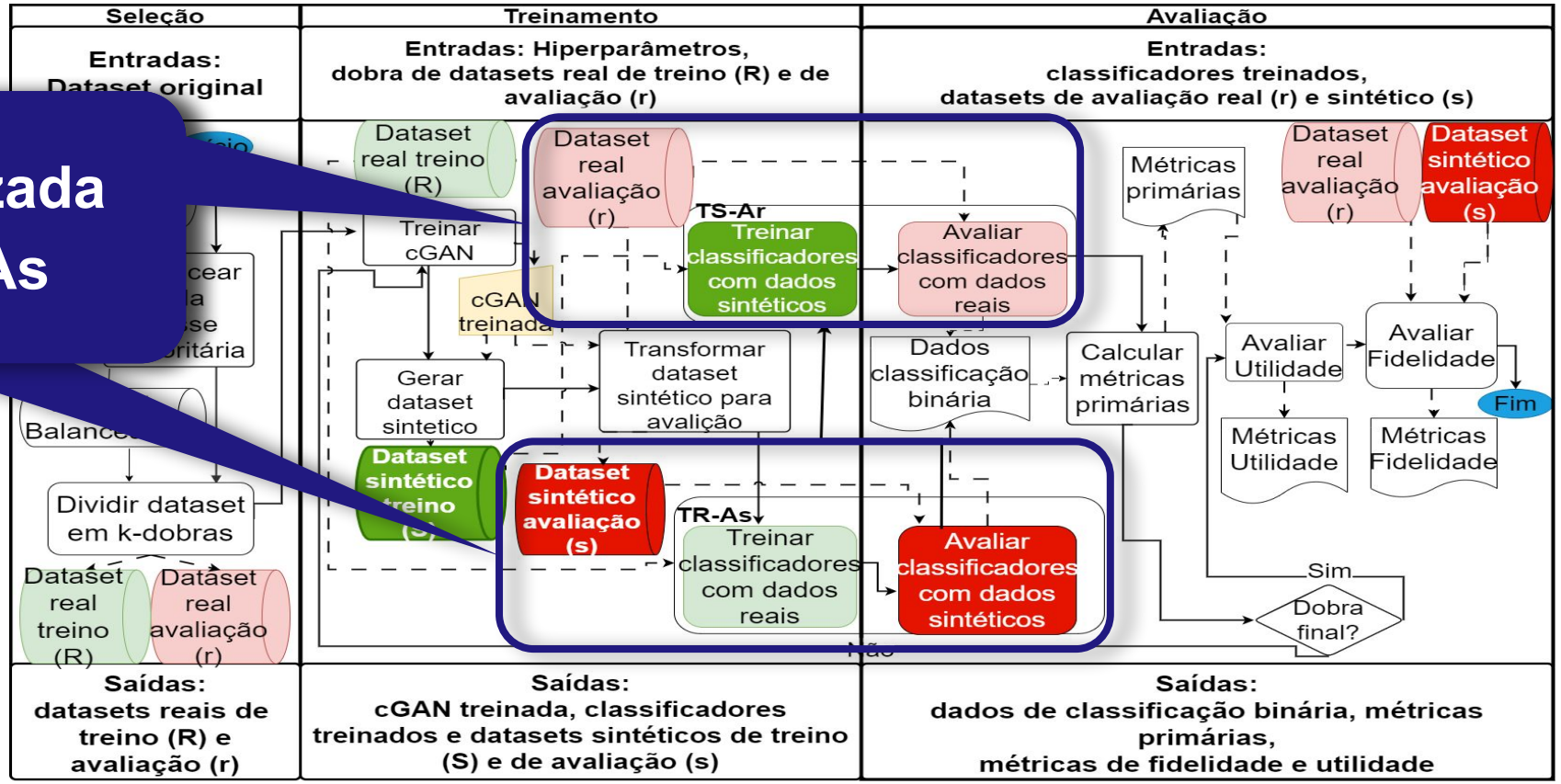
Datasets (real e sintético) de treino e avaliação



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
VIII Salão de Ferramentas do SBSeg 2024
https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/30126

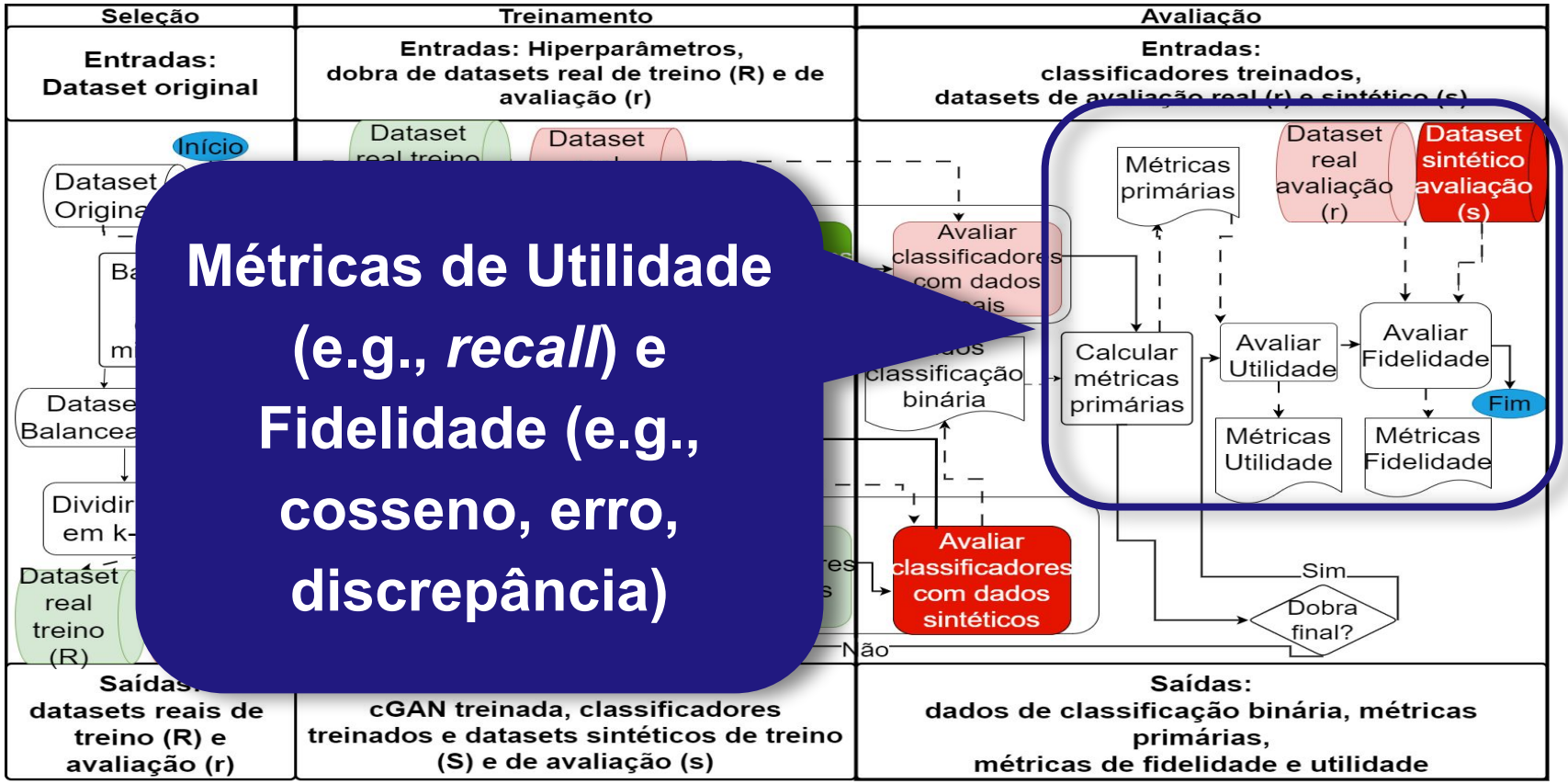
MalSynGen

Avaliação cruzada
TS-Ar e TR-As



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
VIII Salão de Ferramentas do SBSeg 2024
https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/30126

MalSynGen



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
VIII Salão de Ferramentas do SBSeg 2024
https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/30126

MalSynGen

Título: Enabling Data-driven Innovation with Synthetic Data

Vyas Sekar

Carnegie Mellon University

Resumo: Today in computer systems and security research, lack of access to realistic and diverse data from multiple deployments hampers innovation; e.g., products trained on data not representative of environment, there is no way to quantitatively assess products; machine learning workflows experiences data drift, and product audit/feedback is not quantitative. The result today is poo in debugging/reproduction/resolution, and impossibility to share i

In this talk, we will discuss our research outcomes on demonstratin synthetic data using Deep Generative Models (DGMs) to address telemetry, anomaly detection, model training). We have identified privacy challenges and tradeoffs in existing approaches. By synthe advances in machine learning and privacy, we identify design cho will present some of the key results from our work in applying these relevant datasets and use cases.

Mini Bio: Vyas Sekar is the Tan Family Professor of Electrical and Computer Engineering in the ECE Departn Technologist at Rockfish Data, and the Chief Scientist at Conviva.

His research is broadly at the intersection of networks, systems, and security. His work has been recognize Science of Security prize, the Intel Outstanding Researcher Award, the SIGCOMM Test of Time Award, and r

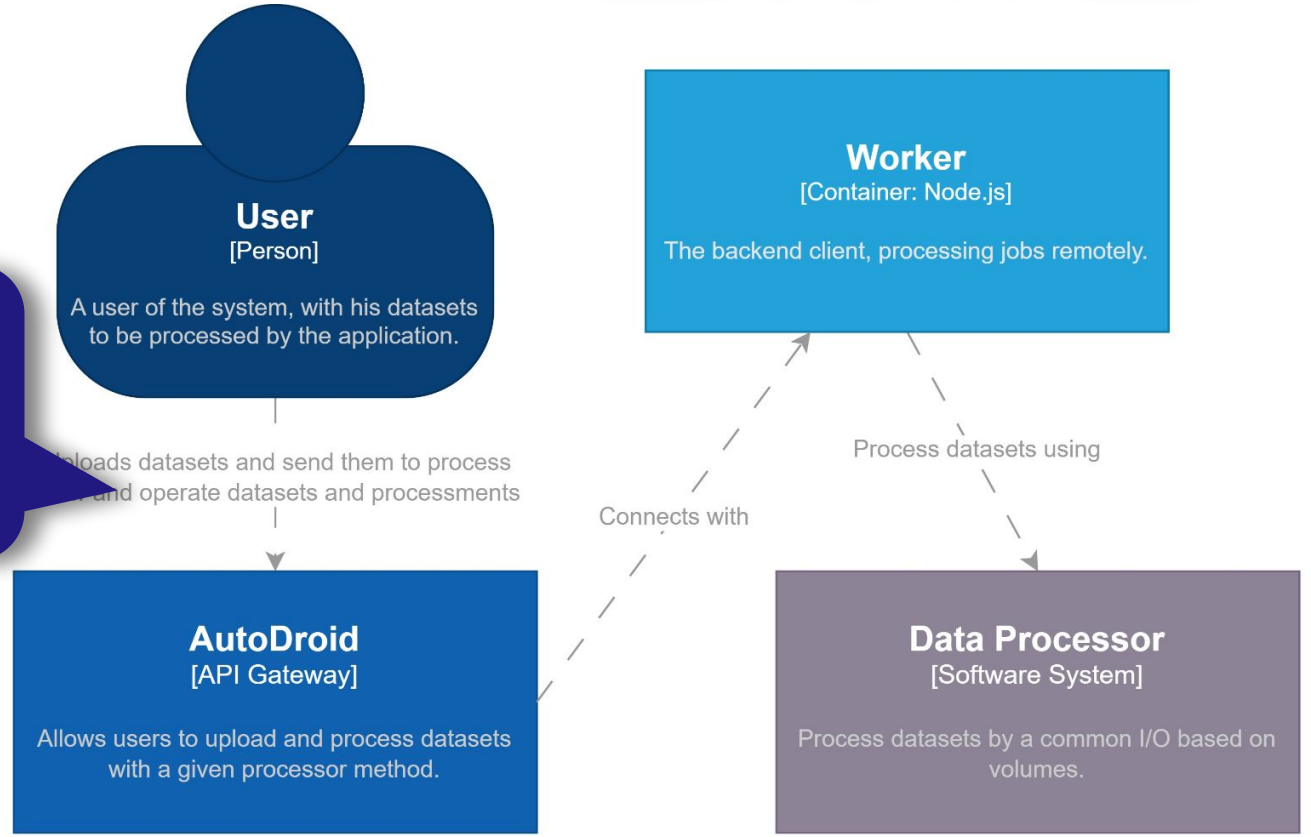
Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. Laboratório de Experimentação
3. MalSynGen (evolução da DroidAugmentor)
4. **Cloud AutoDroid (evolução da AutoDroid)**
5. GUI4MalSynGen (versão minimalista)

Cloud AutoDroid: sistema distribuído

HACKERS DO BEM

API (REST e GraphQL)
Workers distribuídos



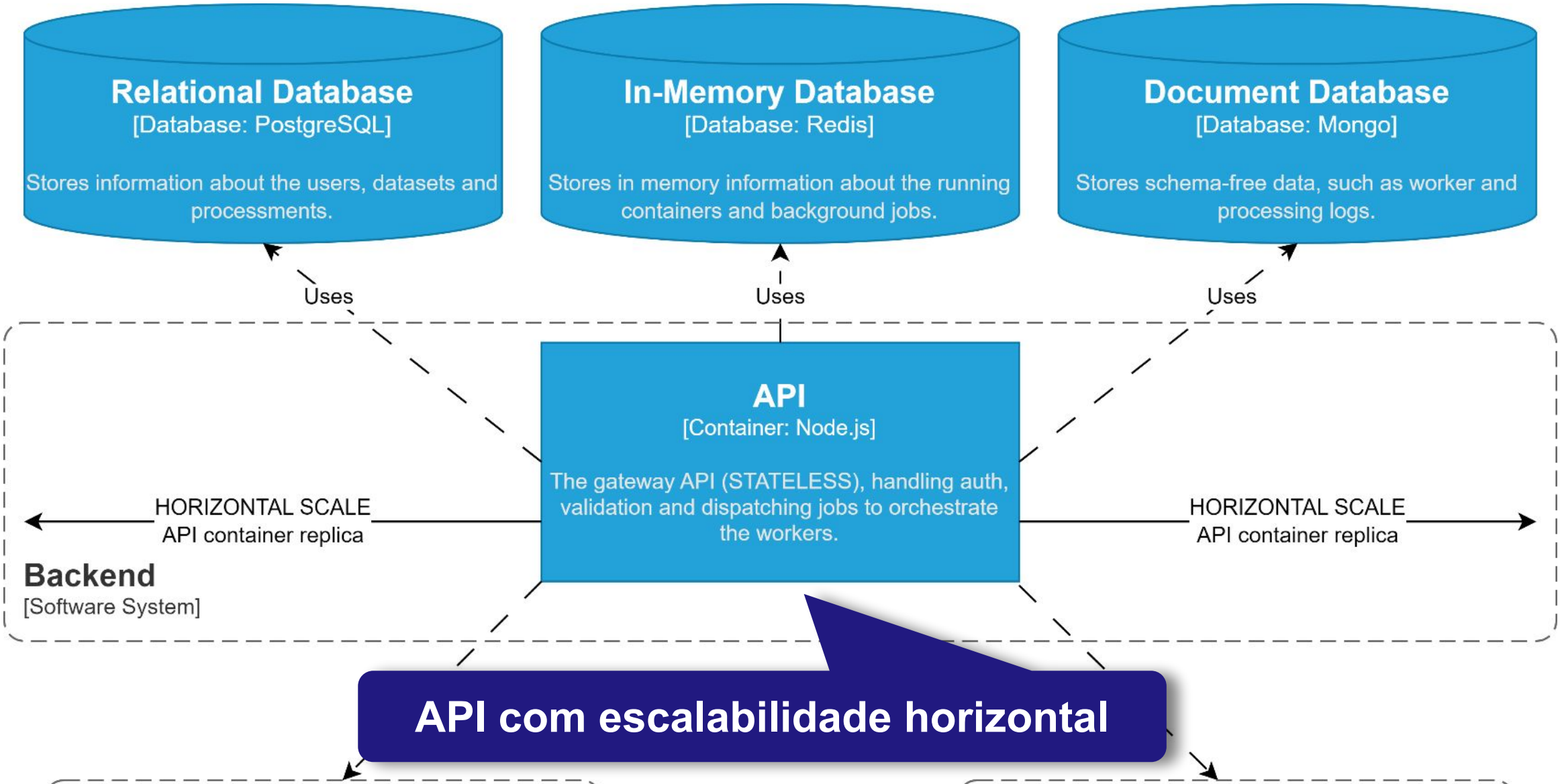
[System Context] AutoDroid
The system context diagram for the Data Processor Gateway System.

Cloud AutoDroid: uma Arquitetura de Backend para Executar Serviços de IA Generativa na Nuvem
Luiz Felipe Laviola, Angelo Gaspar Diniz Nogueira, Diego Kreutz, Rodrigo Brandão Mansilha
VIII Escola Regional de Engenharia de Software (ERES 2024)
https://drive.google.com/file/d/1Q8acpITbFMp9cVdsMpPvpJghMf4lTQqH/view

(paper)

(slides)

Cloud AutoDroid: sistema distribuído



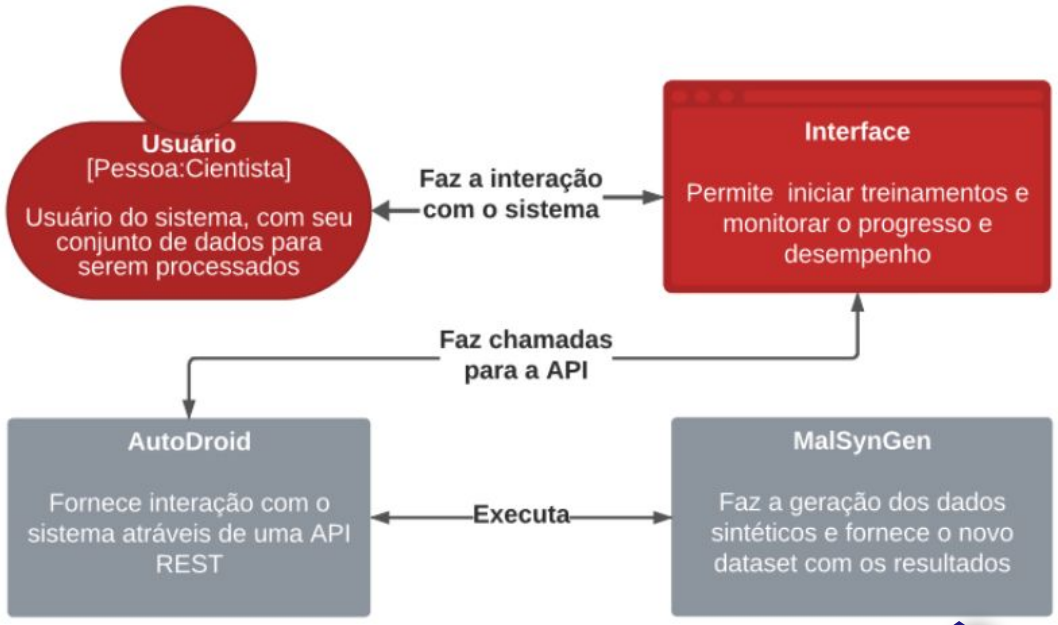
Cloud AutoDroid: limitações atuais

- **Infraestrutura atual**:
  - backend e frontend: 2 máquinas
  - workers (máquinas de processamento): 6 disponíveis
- **Armazenamento**: **cota gratuita** do Firebase Storage limitado a **5GB** e **largura de banda** limitada e **1GB** por dia
- **Auto-escalabilidade**: necessita recursos de cloud e implementação de auto-deploy de workers

Demo Day Malware DataLab

1. Laboratório de Ensino-Aprendizagem
2. Laboratório de Experimentação
3. MalSynGen (evolução da DroidAugmentor)
4. Cloud AutoDroid (evolução da AutoDroid)
5. **GUI4MalSynGen (versão minimalista)**

GUI4MalSynGen



Instância *standalone* para download e teste local

Uma GUI para Hackers do Bem Aprenderem Sobre Malwares Sintéticos
Leonardo Karling Sonco, Angelo Nogueira, Diego Kreutz, Rodrigo Brandão Mansilha
VIII Escola Regional de Engenharia de Software (ERES 2024)
https://github.com/MalwareDataLab/GUI4MalSynGen

(paper)

(GitHub)

Demo Day Malware DataLab

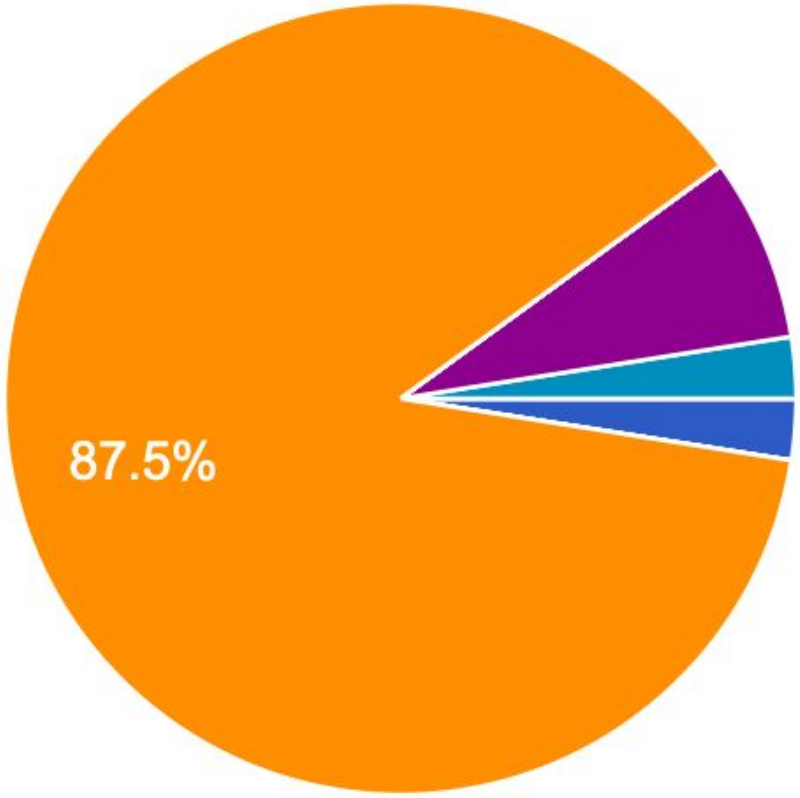
1. Testes com Usuários
2. Prospecção de Mercado
3. Publicações
4. Premiações
5. Direções futuras

Testes com Usuários

27 de novembro de 2024:

- **59 usuários registrados**
- **229 experimentos executados**

Testes com Usuários



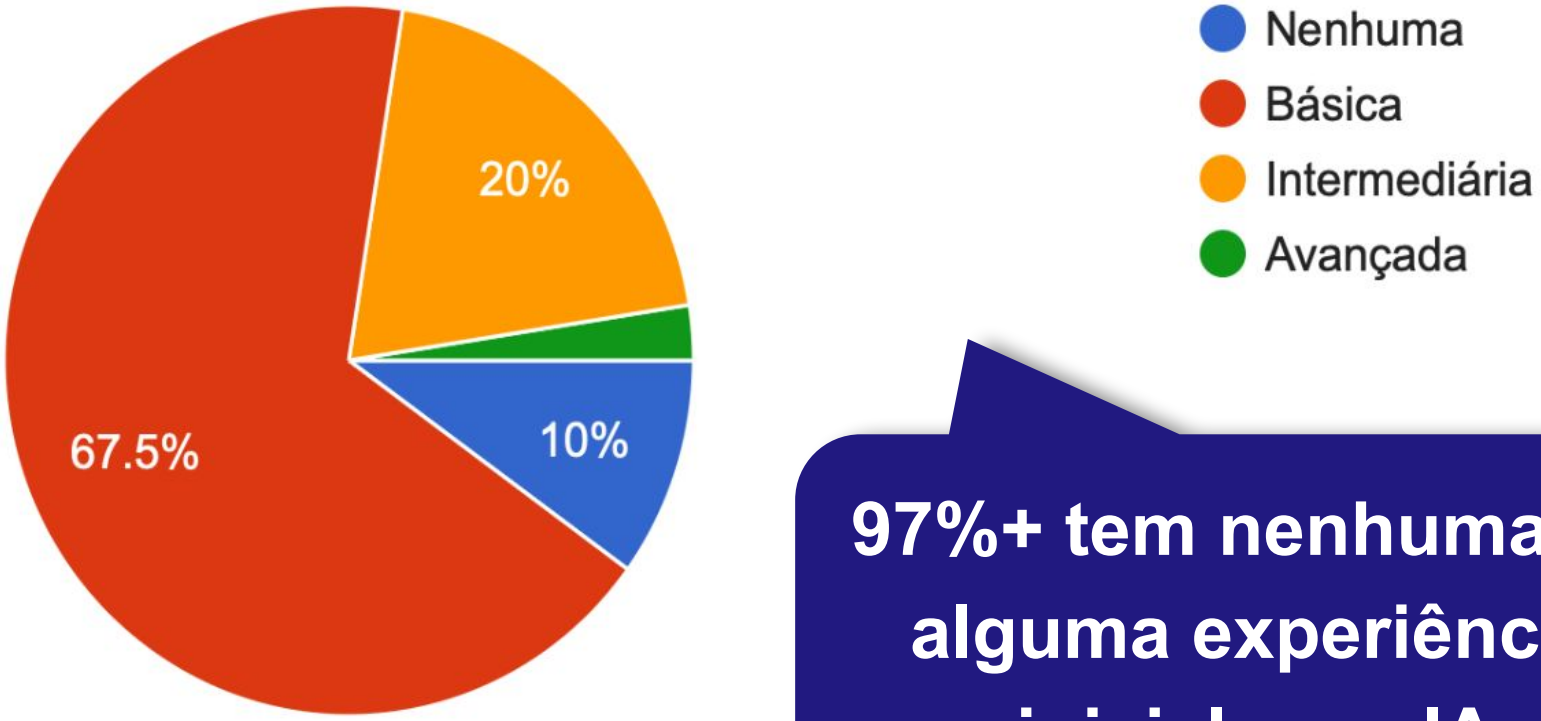
- Ensino Médio
- Ensino Técnico
- Graduação (Bacharelado/Licenciatura)
- Pós-Graduação (Especialização)
- Mestrado
- Doutorado
- Profissional (Atuação no mercado)

4 perfis
40 participantes
4 instituições

Testes com Usuários

Qual é a sua experiência atual com Inteligência Artificial?

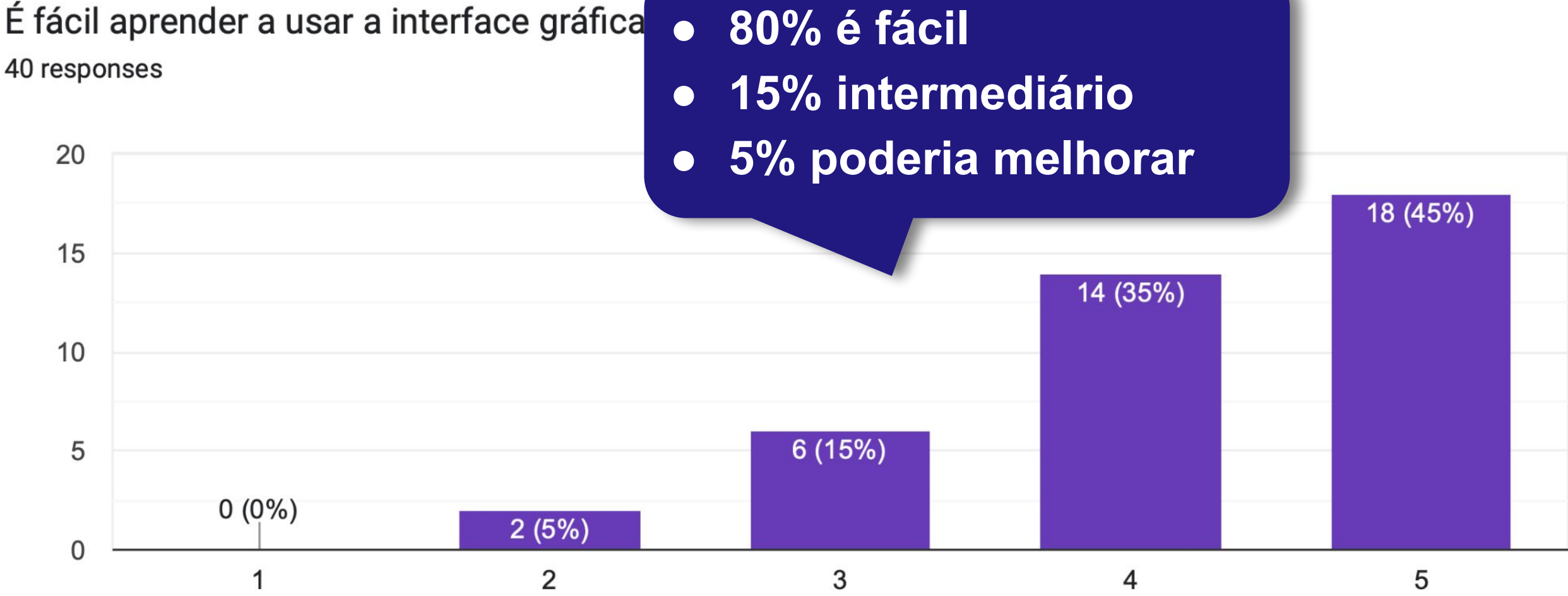
40 responses



97%+ tem nenhuma ou alguma experiência inicial com IA

Testes com Usuários

- 80% é fácil
- 15% intermediário
- 5% poderia melhorar



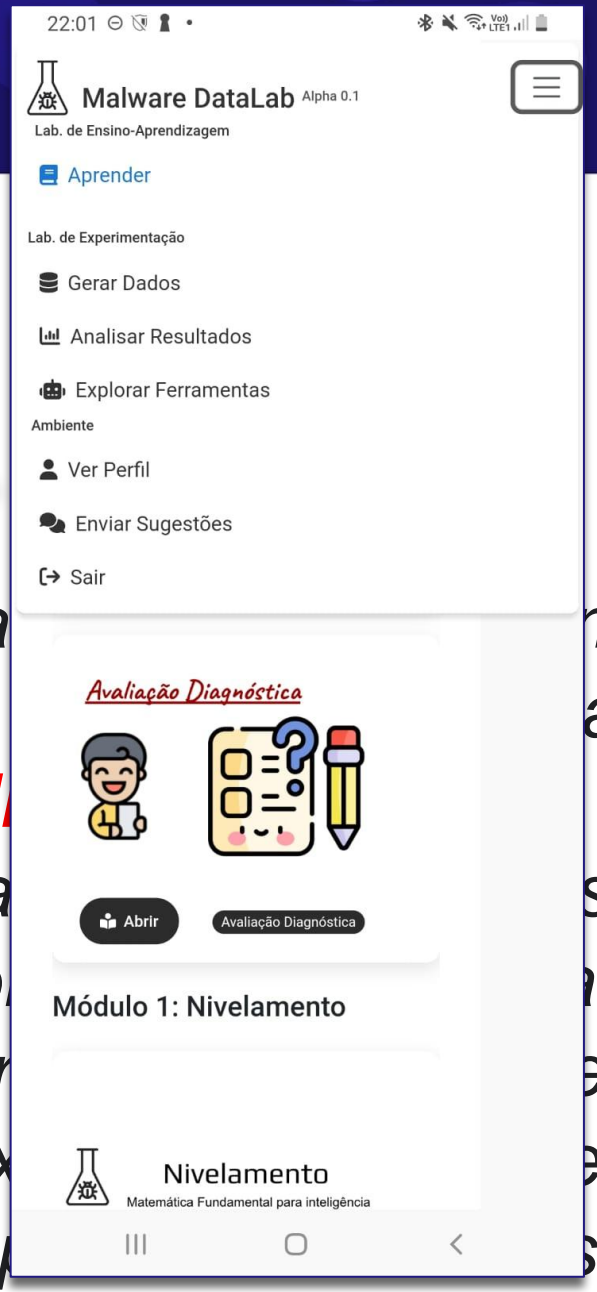
Testes com Usuários

Interface precisa melhorar para versão mobile

"Sobre a questão do uso da plataforma, em alguma modelos de tela dependendo da resolução, os menus não ficam disponíveis da maneira correta. ***Para utilização mobile, o menu não é apresentado corretamente****, apenas a parte dos módulos são listados, a pesquisa funciona, porém se eu precisar seguir o passo a passo dentro do mobile para acessar por exemplo: Gerar Dados, Analisar Resultados, Explorar algoritmo e perfil não são acessíveis por dispositivos menores."*

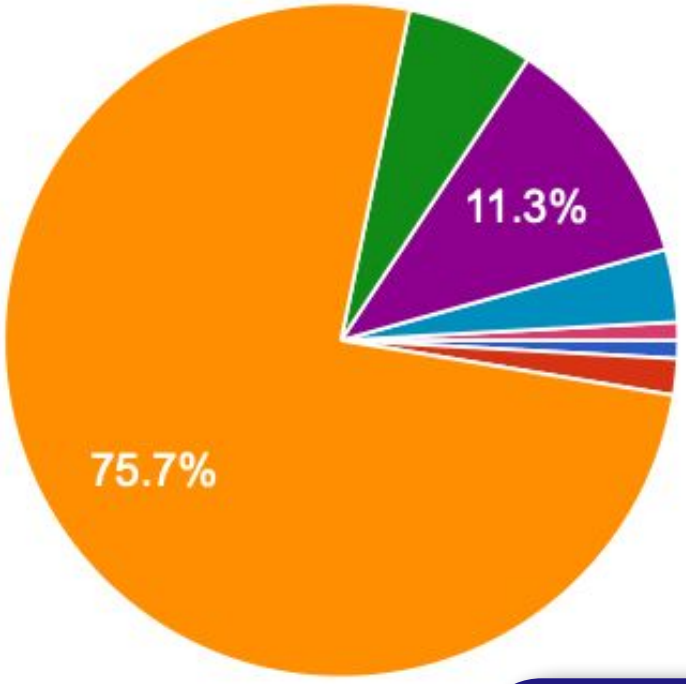
Testes com Usuários

Já estamos trabalhando nisso!



"Sobre a questão do uso da pla ma modelos de tela dependendo da resolução, am disponíveis da maneira correta. ***Para utili menu não é apresentado corretamente****, a s módulos são listados, a pesquisa funciona, po r seguir o passo a passo dentro do mobile par emplo: Gerar Dados, Analisar Resultados, Ex e perfil não são acessíveis por dis s."*

Prospecção de Mercado



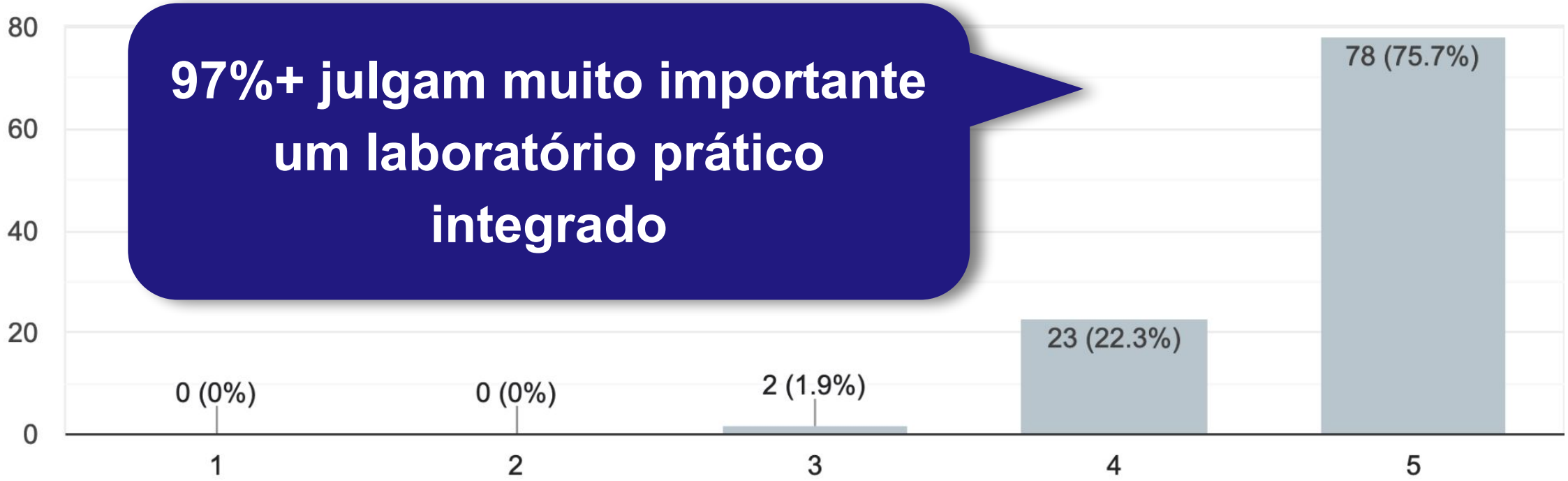
- Ensino Médio
- Ensino Técnico
- Graduação (Bacharelado/ Licenciatura)
- Pós-Graduação (Especialização)
- Mestrado
- Doutorado
- Profissional (Atuação no mercado)

115 participantes de sul a norte do Brasil
96% tem interesse em formação em IA aplicada a cibersegurança

Prospecção de Mercado

Na sua opinião, qual a importância do laboratório prático integrado ao conteúdo de aprendizagem para aprofundar seus conhecimentos e colocar em prática conceitos inovadores de IA?

103 responses



97%+ julgam muito importante um laboratório prático integrado

Publicações

Publicações em eventos em 2024: 7

- **32nd Euromicro** (internacional): 1
- **XXIV SBSeg** (nacional): 2
- **VIII ERES** (regional): 3
- **VIII WRSeg** (regional): 1

Capítulos de livro / minicursos em 2024: 2

1. **DinD-Bench: Impacto de Contêineres Docker em Docker para a Programação Paralela (ERAD)**

Claudio Schepke, Felipe Fava, Diego Kreutz

DOI: https://doi.org/10.5753/sbc.13779.0.3

2. **Uma Introdução sobre Redes Adversárias Generativas (GAN) e suas Aplicações na Cibersegurança (capítulo em ebook de Cibersegurança do UniHacker)**

Karina, Angelo, Anna, Kayua, Diego, Rodrigo, Hendrio

SBSeg24
SÃO JOSÉ DOS CAMPOS

Certificamos que o trabalho **"Geração de dados sintéticos tabulares para detecção de Malwares Android: um estudo de caso",** de autoria de Angelo Diniz (UNIPAMPA), Kayuã Paim (UFRGS), Rodrigo Mansilha (UNIPAMPA) e Diego Kreutz (UNIPAMPA) recebeu o prêmio de

MENÇÃO HONROSA ARTIGO CURTO

na Trilha Principal - Sistemas do 24º Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais.

Lourenço Alves Pereira Jr., Ph.D
Coordenador Geral
Instituto Tecnológico de Aeronáutica

Diego Kreutz, Ph.D
Coordenador Geral
Universidade Federal do Pampa

Trilha Principal - Sistemas

SBSeg24
SÃO JOSÉ DOS CAMPOS

SBC

HACKERS DO BEM

Certificamos que o trabalho **"MalSynGen: redes neurais artificiais na geração de dados tabulares sintéticos para detecção de malware"**, de autoria de Angelo Gaspar Diniz Nogueira (UNIPAMPA), Kayua Oleques Paim (UFAM), Rodrigo Mansilha (UNIPAMPA), Hendrio Luis Bragança (UFRGS) e Diego Kreutz (UNIPAMPA) recebeu o prêmio de

ARTEFATO DESTAQUE

no Salão de Ferramentas (SF) do 24º Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais.

Lourenço Alves Pereira Jr., Ph.D
Coordenador Geral
Instituto Tecnológico de Aeronáutica

Diego Kreutz, Ph.D
Coordenador Geral
Universidade Federal do Pampa

Premiações (3)

Revisores Destaque:

- **SBSeg24** / **CTA**: Angelo Diniz, Diego Kreutz
- **SBSeg24** / **WTICG**: Diego Kreutz, Rodrigo Mansilha
- **SBSeg24** / **SF**: Diego Kreutz, Rodrigo Mansilha

Concorrendo ao *Best Paper* do WRSeg 2024

Otimização de hiperparâmetros da DroidAugmentor para geração de dados sintéticos de malware Android

Angelo Nogueira, Lucas Oliveira, Anna Luiza Gomes, Diego Kreutz, Rodrigo Mansilha

VIII Workshop Regional de Segurança da Informação e de Sistemas Computacionais / ERRC 2024

Direções futuras: Interface

Planejado para Dezembro de 2024:

- **Ajustes finos na interface** a partir do feedback dos usuários
- **Análise de UI e UX** com outros colaboradores experientes
  - **Dr. Williamson Alison Freitas Silva (UNIPAMPA)**
    - http://lattes.cnpq.br/7511013446343990
  - **Dr. Jean Felipe Cheiran (UNIPAMPA)**
    - http://lattes.cnpq.br/9076894715365563

Direções futuras: Interface

- Apoio ao **aprendizado**:
  - Aperfeiçoar o **layout**
  - Revisar e ampliar as **instruções**
  - Adicionar **métricas** (e.g. consumo de recursos)
  - Integração com **Mlflow** (e.g., gráficos dinâmicos e personalizáveis)
- Revisar e melhorar a **usabilidade** geral
- Ajustes de **desempenho**
- **Compatibilidade** com múltiplos tipos de telas
- **Integração** com serviços de *chatbot*

Direções futuras: Cloud AutoDroid

- **Auto-escalabilidade** utilizando computação em Nuvem
- **Sistema de notificação** (e-mail, SMS, WhatsApp, push, etc.)
- **Dashboard administrativo** com estatísticas em tempo real
- **Monitoramento** (e.g., **Mlflow**) dos trabalhos sendo executados
- Estimativas/**cálculos dos custos** de processamento
- **Documentação** via OpenAPI

Direções futuras: MalSynGen

- **Integração** com serviços de *chatbot*
- Generalização da ferramenta **para outros conjuntos de dados**
- **Resumo** dos logs de saída
- Inclusão de mais **métricas de avaliação**
- **Comparação** com outras ferramentas
- **Generalização da ferramenta** para outros conjuntos de dados

Demo Day Malware DataLab

- Papers
- Slides (apresentações)
- Histórico das evoluções
- Vídeos técnicos

https://drive.google.com/drive/folders/1gHqODJZ7AlOGIl6hWm290GbmrGQTAZ0t?usp=sharing

HACKERS DO BEM

Obrigado!

GT Malware DataLab

https://mdl.unihacker.club

Demo Day

29 de novembro de 2024

HACKERS DO BEM_

Video

(overview da plataforma)