

# **Ferramenta de Detecção de ► Phishing Usando Aprendizado de Máquina**

Marina Rijo de Oliveira      RA: 191025501  
Profº Dr. Kelton Augusto Pontara da Costa

# ► TABLE OF CONTENTS

**01**

**INTRODUÇÃO**

**02**

**FUNDAMENTAÇÃO**

**03**

**METODOLOGIA**

**04**

**RESULTADOS**

**05**

**CONCLUSÃO**

**06**

**REFERÊNCIAS**



# ► INTRODUÇÃO

01

# ► INTRODUÇÃO



## RESUMO

Detectar *phishing* é uma das grandes dificuldades da área da segurança da informação na atualidade. Associar aprendizado de máquina à uma interface gráfica simplificada pode ser um grande passo em direção à um avanço nessa frente.



## OBJETIVO

Desenvolver uma aplicação com interface gráfica simples que utiliza algoritmos de aprendizado de máquina encadeados como mecanismo de detecção de *links* de *phishing*.



# ► FUNDAMENTAÇÃO O 02

# ► PHISHING

## DEFINIÇÃO

*Phishing* é um conceito definido como o uso de técnicas de engenharia social, associada à clonagem da identidade visual de páginas confiáveis para captar dados sensíveis das vítimas.

## DIFICULDADES

Essa modalidade de crime se aproveita do crescente alcance da tecnologia, e da falta de conhecimento e familiaridade das pessoas com a internet, para extrair informações sensíveis como login, senha, dados bancários, etc.

# ▶ APRENDIZADO DE MÁQUINA

## DEFINIÇÃO

Uma área das Ciências da Computação que objetiva o desenvolvimento de sistemas que conseguem aprender a identificar padrões e tomar decisões.

## APRENDIZADO SUPERVISIONADO

É uma vertente do aprendizado de máquina, definida como algoritmos que recebem dados já classificados para treinar modelos, com o intuito de identificar e categorizar novas entradas.

# ▶ ALGORITMOS UTILIZADOS

## ÁRVORE DE DECISÃO

Utiliza-se de uma série de questionamentos e validações hierárquicas para classificar novos dados.

## XGBOOST

Modelo do tipo *ensemble* que utiliza *boosting* por gradiente para combinar modelos, corrigindo os erros obtidos a cada iteração.

## FLORESTA ALEATÓRIA

Modelo do tipo *ensemble*, utiliza-se de diversas árvores de decisão diferentes associadas para obter resultados mais precisos.

## CATBOOST

Semelhante ao XGBoost, porém otimizado para o uso de variáveis categóricas.

## EXTRA TREES

Uma variação do floresta aleatória, utiliza o conjunto de dados inteiro e aleatoriza os nós das sub-árvores.

## REGRESSÃO LOGÍSTICA

É um modelo estatístico linear muito utilizado para problemas de classificação binária.



# ► AVALIAÇÃO DOS MODELOS

## **TP** VERDADEIRO POSITIVO

A entrada é verdadeira e é classificada como verdadeira.

## **TN** VERDADEIRO NEGATIVO

A entrada é falsa e é classificada como falsa.

## **FP** FALSO POSITIVO

A entrada é falsa e é classificada como verdadeira.

## **FN** FALSO NEGATIVO

A entrada é verdadeira e é classificada como falsa.

# ► MÉTRICAS

## ACURÁCIA

Define a porcentagem de resultados classificados corretamente.

$$\text{Acurácia} = \frac{TP + TN}{TP + FP + TN + FN}$$

## PRECISÃO

Define a porcentagem de resultados do tipo verdadeiro positivo.

$$\text{Precisão} = \frac{TP}{TP + FP}$$

# ► MÉTRICAS

## MATRIZ DE CONFUSÃO

Uma matriz de duas dimensões que representa os resultados da classificação binária de um modelo.

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	TP	FP
	NEGATIVO	FN	TN



# ▶ **METODOLOGIA**

**03**

# ► FERRAMENTAS

## PYTHON

Linguagem de programação interpretada, muito utilizada na área de ciência de dados. Facilita o desenvolvimento por ser bem simples e possuir diversas bibliotecas voltadas para a área

## SCIKIT-LEARN

## PANDAS

## MATPLOTLIB

## XGBOOST

## CATBOOST

# ► FERRAMENTAS

## GOOGLE COLAB

Ambiente de programação disponibilizado pela Google, possibilitando o desenvolvimento de scripts de maneira remota.

Foi utilizado para o treinamento e avaliação dos modelos.

## STREAMLIT

*Framework* em Python que facilita o desenvolvimento de páginas web.

Auxilia na criação da interface gráfica de forma simples, abstraindo conceitos de HTML e CSS.

## ► CONJUNTO DE DADOS

Foram seleccionados dois conjuntos de dados durante o projeto, um elaborado por Vrbančič, Fister e Podgorelec em 2020 e outro elaborado por Prasad e Chandra em 2024.

Devido à natureza do projeto e à diversas dificuldades com o *dataset* mais novo, o desenvolvimento foi feito com base no elaborado em 2020.

Este conjunto é composto por 88.647 entradas, divididas da seguinte forma:

**LEGÍTIMAS**

58.000

**PHISHING**

30.647

# ► CONJUNTO DE DADOS

O conjunto original possui 112 características, sendo uma delas a classificação em relação à legitimidade da página em questão.

Foram selecionadas 41 dessas características para o treinamento dos modelos, sendo divididas entre atributos baseados no endereço completo e no domínio do endereço

## ENDEREÇO COMPLETO

Características obtidas a partir do *link* completo.

## DOMÍNIO DO ENDEREÇO

Características obtidas exclusivamente a partir do domínio do *link*.



# ▶ CLASSIFICAÇÃO DOS MODELOS

O *dataset* foi dividido em 70% das entradas para treinamento e 30% para testes.

Os modelos foram classificados e ordenados com base na acurácia, precisão e tempo de execução médio.

**ÁRVORE DE  
DECISÃO**



**XGBOOST**



**REGRESSÃO  
LÓGICA**



**FLORESTA  
ALEATÓRIA**



**EXTRA TREES**

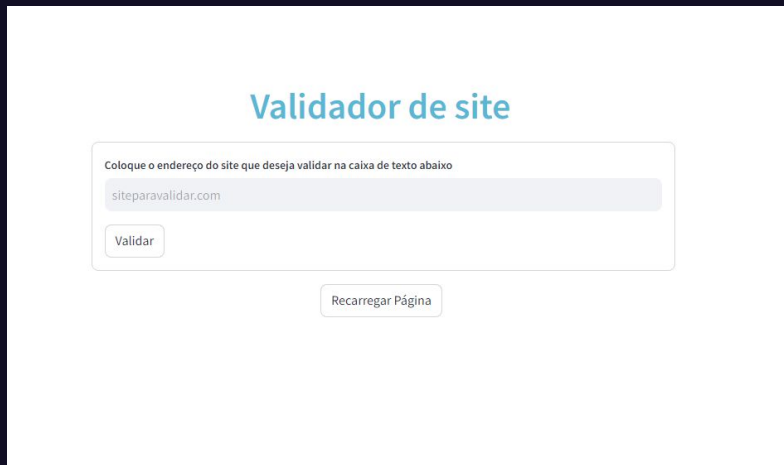


**CATBOOST**

# ► DESENVOLVIMENTO DA APLICAÇÃO

A aplicação recebe um endereço web a ser validado, extrai as informações necessárias para obter uma classificação dos modelos e executa os algoritmos de forma orquestrada, a fim de obter uma maior eficiência, acionando apenas os modelos necessários.

## TELA INICIAL



Validador de site

Coloque o endereço do site que deseja validar na caixa de texto abaixo

siteparavalidar.com

Validar

Recarregar Página

# ► DESENVOLVIMENTO DA APLICAÇÃO

## TELA DE PÁGINA SEGURA

### Validador de site

Site Seguro

O site foi validado e pode ser acessado sem riscos à segurança

Recarregar Página

# ► DESENVOLVIMENTO DA APLICAÇÃO

## TELA DE PÁGINA DUVIDOSA



# ▶ DESENVOLVIMENTO DA APLICAÇÃO

## TELA DE PÁGINA PERIGOSA

### Validador de site

Site Perigoso

O site foi avaliado e não deve ser acessado por apresentar riscos altíssimos à segurança

Recarregar Página



# ► RESULTADOS

04

# ▶ ÁRVORE DE DECISÃO

## RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
92,42%	92,53%	1,1 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	7747	1399
	NEGATIVO	641	16808

# ► XGBOOST

## RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
92,92%	90,85%	1,63 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	8077	1069
	NEGATIVO	813	16636



# ► REGRESSÃO LOGÍSTICA

## RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
89,39%	88,63%	5,91 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	7256	1890
	NEGATIVO	913	16518

# ► FLORESTA ALEATÓRIA

## RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
93,45%	91,99%	6,77 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	8114	1032
	NEGATIVO	705	16744

## ▶ EXTRA TREES

### RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
93,41%	93,87%	8,35 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	8059	1087
	NEGATIVO	676	16773

# ► CATBOOST

## RESULTADOS OBTIDOS

ACURÁCIA	PRECISÃO	EXECUÇÃO
93,32%	91,63%	22,7 segundos

		Classe Prevista	
		POSITIVO	NEGATIVO
Classe Verdadeira	POSITIVO	8110	1036
	NEGATIVO	741	16708

## ► RESULTADOS GERAIS

	ACURÁCIA	PRECISÃO	EXECUÇÃO
ÁRVORE DE DECISÃO	92,42%	92,53%	1,1 segundos
XGBOOST	92,92%	90,85%	1,63 segundos
REGRESSÃO LOGÍSTICA	89,39%	88,63%	5,91 segundos
FLORESTA ALEATÓRIA	93,45%	91,99%	6,77 segundos
EXTRA TREES	93,41%	93,87%	8,35 segundos
CATBOOST	93,32%	91,63%	22,7 segundos



# ► CONCLUSÃO

05

# ► CONCLUSÃO

## OBJETIVOS ALCANÇADOS

Este projeto objetiva o desenvolvimento de uma aplicação que auxilia na prevenção de ataques de *phishing*. A aplicação foi construída e seguiu sua principal diretiva, ser simples de usar.

## DIFICULDADES ENCONTRADAS

Devido às alterações necessárias ao conjunto de dados, os resultados da aplicação não são tão acurados e precisos quanto os resultados obtidos nos testes.

## TRABALHOS FUTUROS

Futuros trabalhos podem construir um *dataset* mais robusto dedicado a essa abordagem, melhorando a precisão e acurácia dos modelos.



# ► **REFERÊNCIAS**



**06**





## ► REFERÊNCIAS

- ALEROUD, A.; ZHOU, L. Phishing environments, techniques, and countermeasures: A survey. Computers & Security, Elsevier, v. 68, p. 160-196, 2017.
- ALKHALIL, Z.; HEWAGE, C.; NAWAF, L.; KHAN, I. Phishing attacks: A recent comprehensive study and a new anatomy. Frontiers in Computer Science, Frontiers Media SA, v. 3, p. 563060, 2021.
- COUTINHO, V. M. Detecção de páginas de phishing utilizando aprendizado de máquina. 2023.
- FAWCETT, T. An introduction to roc analysis. Pattern recognition letters, Elsevier, v. 27, n. 8, p. 861-874, 2006.
- IBRAHIM, A. A.; RIDWAN, R. L.; MUHAMMED, M. M.; ABDULAZIZ, R. O.; SAHEED, G. A. Comparison of the catboost classifier with other machine learning methods. International Journal of Advanced Computer Science and Applications, v. 11, 2020. Disponível em: <<https://api.semanticscholar.org/CorpusID:232846952>>.
- KRAMER, O.; KRAMER, O. Scikit-learn. Machine learning for evolution strategies, Springer, p. 45-53, 2016.
- MCKINNEY, W. et al. pandas: a foundational python library for data analysis and statistics. Python for high performance and scientific computing, Seattle, v. 14, n. 9, p. 1-9, 2011.
- MONTAGNER, A. S.; WESTPHALL, C. M. Uma breve análise sobre phishing. Revista ComInG-Communications and Innovations Gazette, v. 6, n. 1, p. 46-56, 2022.

## ► REFERÊNCIAS

- MÜLLER, A. C.; GUIDO, S. Introduction to machine learning with Python: a guide for data scientists. [S.l.]: "O'Reilly Media, Inc.", 2016.
- PRASAD, A.; CHANDRA, S. PhiUSIIL Phishing URL (Website). 2024. UCI Machine Learning Repository. DOI: <https://doi.org/10.1016/j.cose.2023.103545>.
- RESNICK, N. E.; BASTOS-FILHO, C. J. A. Aplicação de aprendizado de máquinas para detecção de urls phishing. Revista de Engenharia e Pesquisa Aplicada, v. 9, n. 1, p. 41-49, 2024.
- SOUZA, J. A.; MASCARENHAS, D. M. Detecção de ataques de phishing em tempo real utilizando algoritmos de aprendizado de máquina. In: SBC. Anais Estendidos do XXIII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais. [S.l.], 2023. p. 165-176.
- STREAMLIT. Streamlit documentation. 2024. Disponível em: <<https://docs.streamlit.io/>>. Acesso em: 3 abr. 2024.
- VRBANČIČ, G.; FISTER, I.; PODGORELEC, V. Datasets for phishing websites detection. Data in Brief, v. 33, p. 106438, 2020. ISSN 2352-3409. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352340920313202>>.