A UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO COMBATE AO *PHISHING*
THE USE OF ARTIFICIAL INTELLIGENCE IN COMBATING PHISHING

Thayná Marostica Machado da Silva[1]
Rafael de Sá Mascarenhas[2]

RESUMO: Este artigo analisa o fenômeno do *phishing* sob a ótica da Engenharia Social, investigando a manipulação de vieses cognitivos para a exfiltração de dados sensíveis. Por meio de uma pesquisa bibliográfica, examina-se o papel da Inteligência Artificial (IA) na identificação e mitigação de ameaças cibernéticas. A pesquisa estratifica a eficácia de algoritmos de Machine Learning (ML), como Support Vector Machine (SVM) e XGBoost, na análise estatística e reconhecimento de padrões em metadados. Adicionalmente, explora-se a aplicação de redes neurais de Longa Memória de Curto Prazo (LSTM) e Processamento de Linguagem Natural (PLN) no domínio do Deep Learning (DL), visando a detecção de gatilhos semânticos de urgência e coerção em comunicações eletrônicas. Os resultados indicam que a integração híbrida de modelos de ML e DL potencializa a acurácia preditiva, superando limitações de abordagens isoladas. Discute-se, ainda, a natureza dual da IA evidenciando seu uso ambivalente: como vetor de defesa robusta e como ferramenta de automação para ataques de *spear-phishing* em larga escala. Conclui-se que a implementação de sistemas multicamadas baseadas em IA é imperativa para a resiliência cibernética de usuários em diversos níveis de proficiência técnica.
Palavras-chave: Cibersegurança; Deep Learning; Inteligência Artificial; Machine Learning; Prevenção.

ABSTRACT: This paper analyzes the phenomenon of phishing through the lens of Social Engineering, investigating the manipulation of cognitive biases for sensitive data exfiltration. Through a systematic literature review, the role of Artificial Intelligence (AI) in identifying and mitigating cyber threats is examined. The research stratifies the efficacy of Machine Learning (ML) algorithms, such as Support Vector Machine (SVM) and XGBoost, in statistical analysis and pattern recognition within metadata. Additionally, it explores the application of Long Short-Term Memory (LSTM) networks and Natural Language Processing (NLP) within the Deep Learning (DL) domain, aiming to detect semantic triggers of urgency and coercion in electronic communications. The results indicate that the hybrid integration of ML and DL models enhances predictive accuracy, overcoming the limitations of isolated approaches. Furthermore, the dual nature of AI is discussed, highlighting its ambivalent use: as a robust defense vector and as an automation tool for large-scale spear-phishing attacks. It is concluded that the implementation of AI-based multi-layered systems is imperative for the cyber resilience of users across diverse levels of technical proficiency.
Keywords: Artificial Inteligence; Cybersecurity; Deep Learning; Machine Learning; Prevention.

Mestranda em Ciências da Computação - Universidade Federal de São Carlos - E-mail: thaynamms@estudante.ufscar.br [1]
Mestre em Computação Aplicada – Universidade Presbiteriana Mackenzie - E-mail: rafael.mascarenhas@mackenzie.br [2]

1 INTRODUÇÃO

O *Phishing*, embora historicamente consolidado, permanece como uma das táticas de exfiltração de dados mais persistentes e eficazes no ecossistema de ameaças cibernéticas contemporâneo. Fundamentado em vetores de Engenharia Social, o ataque visa manipular psicologicamente a vítima, explorando suas vulnerabilidades, para que informe seus dados sensíveis de forma intencional (Oliveira *et al.*, 2025). No cenário brasileiro, a criticidade deste fenômeno é evidenciada por métricas alarmantes: em 2022, registrou-se uma média de 1.540 incidentes semanais, com uma trajetória de crescimento acentuada pela digitalização acelerada de serviços (Junior; Cunha, 2022).

A despeito dos esforços em programas de conscientização, a sofisticação dos ataques, que agora contam com o auxílio de IAs Generativas para criar comunicações personalizadas e convincentes, impõe limites à eficácia de medidas puramente educativas. Diante disso, torna-se imperativo o desenvolvimento de sistemas de detecção automatizados que operem de forma contínua e transparente, mitigando o erro humano sem comprometer a usabilidade do sistema.

A Inteligência Artificial (IA) emerge como o paradigma para a segurança preditiva. Definida pela aplicação de algoritmos de alta complexidade voltados ao processamento massivo de dados e otimização de funções sistêmicas (MORAIS; BRANCO, 2023), a IA permite a identificação de anomalias em *datasets* diversos em origem e formato. A eficácia dessa tecnologia em identificar padrões complexos já é observada em diversos domínios críticos. Na área da segurança da informação, por exemplo, algoritmos de Machine Learning e Deep Learning podem ser aplicados na prevenção e respostas automáticas a detecções suspeitas (Monfre *et al.*, 2023). Essa capacidade de processamento de variáveis multimodais fundamenta a transposição dessas tecnologias para a cibersegurança, onde a complexidade dos ataques de *phishing* exige a mesma acurácia diagnóstica.

O desafio científico, contudo, reside na seleção e no treinamento de subáreas específicas, dado que as variantes de *phishing* apresentam comportamentos distintos, desde a ofuscação de URLs até a manipulação semântica em comunicações de texto. O presente trabalho propõe uma investigação bibliográfica acerca da eficácia da IA na detecção e mitigação de ataques de *phishing*. A análise foca na convergência entre

Machine Learning e modelos de Processamento de Linguagem Natural, examinando como modelos estatísticos e redes neurais profundas processam indicadores de comprometimento (IoCs). Ao integrar a análise técnica da Engenharia Social com o estado da arte em computação cognitiva, este estudo visa sistematizar o conhecimento necessário para a construção de defesas resilientes contragolpes digitais em tempo real.

1.1 OBJETIVO GERAL

O objetivo geral deste estudo é analisar a eficácia das arquiteturas de Machine Learning (Support Vector Machine e XGBoost) e Deep Learning (redes LSTM e modelos de Processamento de Linguagem Natural) na detecção proativa de ataques de *phishing*, avaliando sua capacidade de gerar alertas precisos e mitigar a vulnerabilidade do usuário final frente a vetores de engenharia social.

1.2 OBJETIVOS ESPECÍFICOS

- Sistematizar o estado da arte acerca das variantes de *phishing* e suas táticas de evasão, estabelecendo uma taxonomia dos desdobramentos atuais no cenário de ameaças cibernéticas;
- Avaliar o desempenho de algoritmos de IA (como SVM, XGBoost e Redes Neurais LSTM) na identificação de padrões maliciosos em datasets de URLs e e-mails, comparando métricas de acurácia e falsos positivos;
- Discutir o paradigma da IA Adversária, examinando como criminosos utilizam a automação e modelos generativos para sofisticar a criação e a distribuição de campanhas de *phishing*.

2 METODOLOGIA

A presente pesquisa caracteriza-se como uma pesquisa bibliográfica de natureza qualitativa. Segundo Gerhardt e Silveira (2009), a pesquisa qualitativa não se preocupa com representatividade numérica, mas com o aprofundamento da compreensão de um grupo social ou organização. Para garantir o rigor científico na

seleção do corpus documental, seguiu-se o protocolo de Sampaio (2022), que estabelece etapas estruturadas de busca, filtragem e análise crítica, permitindo uma síntese fiel do estado da arte sobre o combate ao *phishing* via IA.

2.1 ESTRATÉGIA DE BUSCA E PROTOCOLO DE SELEÇÃO

Para o levantamento do corpus documental, foi consultada a base de dados Google Scholar, selecionada por sua abrangência na indexação de periódicos científicos, teses e anais de eventos de relevância acadêmica. Cada publicação identificada foi submetida aos seguintes critérios de elegibilidade:

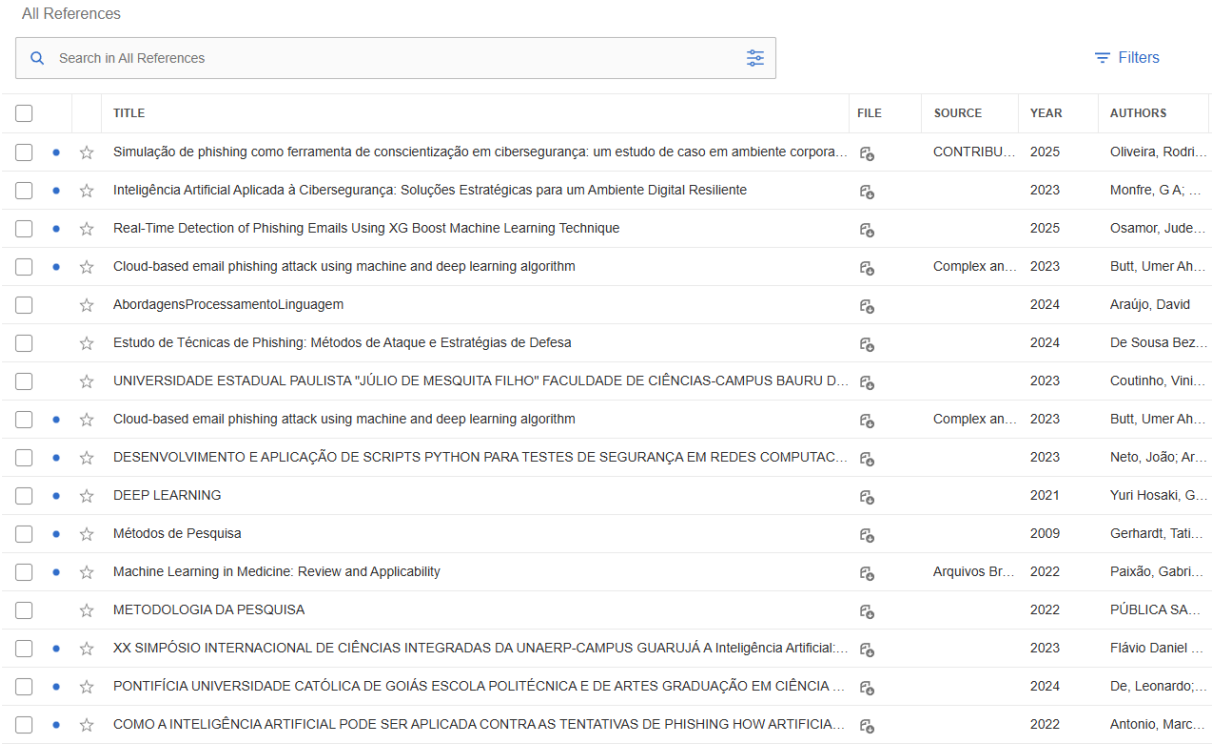
- Pertinência Temática: Verificação da aderência aos eixos centrais: *Phishing*, Engenharia Social e Inteligência Artificial. Publicações que não apresentavam convergência direta com os objetivos da pesquisa foram sumariamente descartadas em favor do tópico subsequente.
- Recorte Temporal: Estabeleceu-se uma janela de dez anos (2020–2025), intervalo que compreende a ascensão das arquiteturas de Redes Neurais Profundas e a crescente sofisticação das ameaças cibernéticas.
- Filtro Linguístico: Restringiu-se a seleção a materiais redigidos em português e inglês, assegurando a precisão na interpretação terminológica e técnica dos algoritmos discutidos.

2.2 GESTÃO E SISTEMATIZAÇÃO DO REFERENCIAL TEÓRICO

Após a etapa de filtragem, o gerenciamento do material selecionado foi operacionalizado através do software Mendeley Reference Manager. A utilização desta ferramenta visou assegurar a integridade dos metadados (autoria, ano e título) e a padronização normativa das citações e referências ao longo do manuscrito.

A organização interna do referencial foi estratificada por meio de uma arquitetura de pastas temáticas e indexação por etiquetas (tags), conforme ilustrado na Figura 1.

Figura 1 - Levantamento das publicações



	TITLE	FILE	SOURCE	YEAR	AUTHORS
●	Simulação de phishing como ferramenta de conscientização em cibersegurança: um estudo de caso em ambiente corpora…		CONTRIBU…	2025	Oliveira, Rodri…
●	Inteligência Artificial Aplicada à Cibersegurança: Soluções Estratégicas para um Ambiente Digital Resiliente			2023	Monfre, G A; …
●	Real-Time Detection of Phishing Emails Using XG Boost Machine Learning Technique			2025	Osamor, Jude…
●	Cloud-based email phishing attack using machine and deep learning algorithm		Complex an…	2023	Butt, Umer Ah…
	AbordagensProcessamentoLinguagem			2024	Araújo, David
	Estudo de Técnicas de Phishing: Métodos de Ataque e Estratégias de Defesa			2024	De Sousa Bez…
	UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO" FACULDADE DE CIÊNCIAS-CAMPUS BAURU D…			2023	Coutinho, Vini…
●	Cloud-based email phishing attack using machine and deep learning algorithm		Complex an…	2023	Butt, Umer Ah…
●	DESENVOLVIMENTO E APLICAÇÃO DE SCRIPTS PYTHON PARA TESTES DE SEGURANÇA EM REDES COMPUTAC…			2023	Neto, João; Ar…
●	DEEP LEARNING			2021	Yuri Hosaki, G…
●	Métodos de Pesquisa			2009	Gerhardt, Tati…
●	Machine Learning in Medicine: Review and Applicability		Arquivos Br…	2022	Paixão, Gabri…
	METODOLOGIA DA PESQUISA			2022	PÚBLICA SA…
●	XX SIMPÓSIO INTERNACIONAL DE CIÊNCIAS INTEGRADAS DA UNAERP-CAMPUS GUARUJÁ A Inteligência Artificial:…			2023	Flávio Daniel …
●	PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS ESCOLA POLITÉCNICA E DE ARTES GRADUAÇÃO EM CIÊNCIA …			2024	De, Leonardo;…
●	COMO A INTELIGÊNCIA ARTIFICIAL PODE SER APLICADA CONTRA AS TENTATIVAS DE PHISHING HOW ARTIFICIA…			2022	Antonio, Marc…

Fonte: Elaborado pelos autores (2026)

Como demonstrado no mapeamento das produções científicas, o corpus documental foi estratificado em quatro eixos distintos para fundamentar a progressão teórica do estudo. O domínio de *Phishing* e Engenharia Social concentra as fontes voltadas aos vetores de ataque e à manipulação comportamental, enquanto o eixo de IA, agrupa o referencial sobre processamento massivo de dados e lógica algorítmica.

Por fim, a categoria de Machine Learning (ML) e Deep Learning (DL) que reúne a literatura técnica sobre modelos preditivos e classificação automatizada de ameaças, ao passo que o segmento de Metodologia de Pesquisa consolida as diretrizes procedimentais que nortearam a pesquisa.

3 REFERENCIAL TEÓRICO

3.1 *PHISHING* E ENGENHARIA SOCIAL

A Engenharia Social consiste em uma estratégia de ataque baseada na manipulação psicológica de indivíduos para a obtenção de dados sensíveis ou o acesso a sistemas restritos, explorando vulnerabilidades do comportamento humano

em detrimento de intrusões estritamente técnicas (Alves, 2024). Para tanto, utiliza-se de gatilhos como o senso de urgência ou a simulação de relações de confiança, de forma que independente do nível de conhecimento técnico e treinamento esses ataques podem ser realizados com sucesso (Oliveira *et al.*, 2025). Dentre as variações de *phishing* associadas a essa prática, o Spear *Phishing* caracteriza-se por comunicações eletrônicas personalizadas para alvos específicos com o intuito de instalar malwares e extrair dados, enquanto o Clone *Phishing* baseia-se na replicação de e-mails legítimos anteriormente enviados, substituindo links originais por endereços maliciosos (Junior; Cunha, 2022).

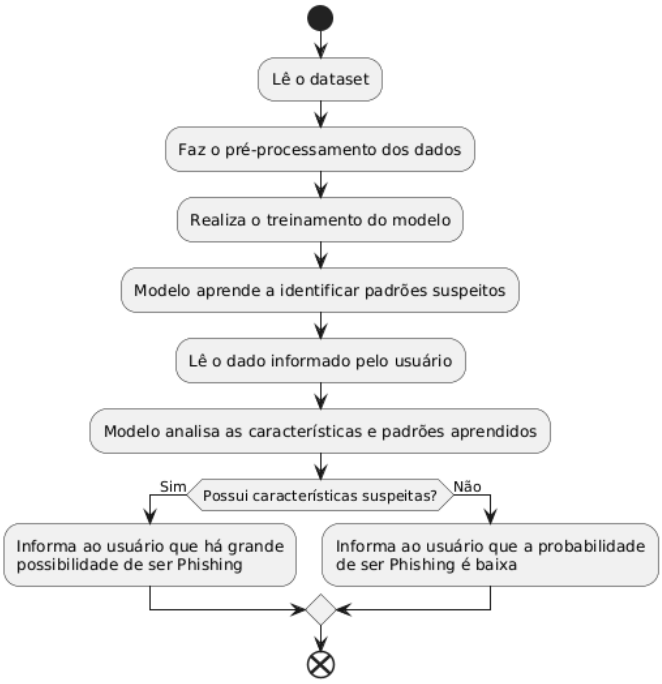
3.2 INTELIGÊNCIA ARTIFICIAL E O EMPREGO DE MACHINE LEARNING

A Inteligência Artificial atua no processamento de volumes de dados que excedem a capacidade de análise manual, permitindo a identificação de padrões de ataque (Morais; Branco, 2023).

Dentro desse domínio, o Machine Learning (ML) integra técnicas matemáticas e estatísticas a algoritmos para identificar padrões em conjuntos de variáveis e realizar predições. Através do treinamento com datasets de volumetria expressiva, os modelos tornam-se capazes de detectar anomalias em elementos como URLs e metadados de comunicações eletrônicas (Coutinho, 2023).

Por fundamentar-se em métodos estatísticos, esses algoritmos quantificam características para calcular a probabilidade de uma ameaça, classificando a natureza da interação e auxiliando na prevenção de ataques de maneira automatizada (Junior; Cunha, 2022). O fluxo operacional desses modelos compreende desde a extração de atributos até a decisão final, fornecendo a estrutura para a mitigação de riscos decorrentes da Engenharia Social, conforme visto na Figura 2, uma exemplificação vislumbrada na literatura.

Figura 2 - Fluxo de funcionamento de algoritmos de ML



Fonte: Elaborado pelos autores (2026)

Como visualizado na Figura 2, o fluxo operacional dos algoritmos de ML inicia-se com a coleta de datasets e o subsequente pré-processamento dos dados, etapa em que as características são normalizadas para o treinamento do modelo. Quando a análise recai sobre textos, esse estágio assume maior profundidade técnica ao envolver métodos como a Tokenização, que segmenta o conteúdo linguístico em unidades menores, e o TF-IDF (Term Frequency-Inverse Document Frequency), utilizado para mensurar a relevância estatística de termos específicos em relação ao corpus documental.

Esse refinamento dos dados brutos é necessário para garantir a eficácia da classificação. Segundo Coutinho (2023, p. 24), "o objetivo de um classificador é, dado um conjunto de dados de entrada, determinar a qual classe estes dados pertencem". Dessa forma, após a extração de atributos e o aprendizado de padrões, o sistema torna-se capaz de analisar novas entradas em tempo real, fornecendo o fundamento necessário para o uso de tecnologias preditivas na mitigação de riscos oriundos da Engenharia Social.

3.3 DEEP LEARNING E PROCESSAMENTO DE LINGUAGEM NATURAL

De acordo com Hosaki e Ribeiro (2021), o Deep Learning (DL) compreende um conjunto de técnicas que utilizam redes neurais artificiais multicamadas para a resolução de problemas por meio de níveis de abstração. No contexto do *Phishing*, o DL é aplicado para a classificação e detecção de ataques, utilizando, por exemplo, arquiteturas de redes neurais para o processamento de sequências, o que permite a identificação de padrões e a categorização da ameaça. Esta abordagem fornece ao usuário informações contextuais sobre a natureza do incidente através de uma análise que considera a complexidade e a profundidade dos dados.

Conforme o estudo de Araújo (2024), o Processamento de Linguagem Natural (PLN) atua como um campo da Inteligência Artificial voltado ao processamento e compreensão da linguagem humana por meio de algoritmos. A integração entre PLN e redes neurais profundas possibilita a análise semântica de textos, permitindo a identificação de indícios de urgência ou coerção. Esta sinergia técnica é o que permite superar as limitações da análise puramente estatística de metadados, auxiliando na sinalização de comunicações suspeitas ao examinar o conteúdo linguístico e fornecendo insights que auxiliam o usuário na percepção de tentativas de manipulação psicológica.

4 RESULTADOS E DISCUSSÃO

A análise dos resultados demonstra que a eficácia na detecção de *phishing* reside na transição da análise de indicadores superficiais para a compreensão contextual. Enquanto os algoritmos de Machine Learning, como o SVM e o XGBoost, operam de forma complementar na identificação de anomalias em metadados e URLs suspeitas, o emprego do Deep Learning resolve a lacuna da análise estritamente técnica ao focar na interpretação semântica das mensagens.

4.1 DESEMPENHO DOS ALGORITMOS DE DETECÇÃO

Foram analisados estudos experimentais que fornecem evidências quantitativas sobre o desempenho de diferentes algoritmos aplicados no contexto de detecção de *Phishing* E-mail.

O primeiro estudo analisado, realizado por Osamor *et al.* (2025), aplicou algoritmos de Random Forest, Árvore de Decisão, XGBoost e Regressão Lógica para

treinarem por meio de um banco de dados públicos com aproximadamente 49.000 e-mails com 80% sendo legítimos e 20% com o resultado conclusivo de serem *phishing*. Foram extraídas 31 features a partir da análise do corpo do e-mail, assunto, cabeçalho, URLs, palavras urgentes e detecção de códigos em HTML.

Já o segundo estudo de Butt *et al*. (2022), focou nos algoritmos de SVM, LSTM e Naive Bayes, treinados pela extração de features com as partes de e-mails categorizados como spam ou não spam proporcionados pelo banco de dados online CSDMC_SPAM.

A tabela 1 apresenta os resultados dos dois estudos, comparando as mesmas métricas utilizadas por ambos.

Tabela 1 – Desempenho dos algoritmos

Algoritmo	Acurácia	Precisão	Recall	F1-Score
Random Forest	98,52%	98%	99%	98,50%
Árvore de Decisão	96,44%	96%	97%	96,50%
XGBoost	98,58%	99%	98%	98,50%
Regressão Lógica	97,05%	97%	97%	97%
SVM	99,62%	99,60%	99,60%	99,60%
LSTM	98%	98%	98%	98%
Naive Bayes	97%	97%	97%	97%

Fonte: Elaborado pelos autores (2026).

Por meio dos resultados é notório que os algoritmos possuem uma taxa alta na detecção de *phishing*, destacando-se o SVM com 99,62% de acurácia. Ao analisar os resultados é perceptível que o XGBoost demonstrou maior equilíbrio entre precisão e recall, essa reciprocidade entre os algoritmos indica um potencial de compatibilidade para serem utilizadas como abordagens híbridas, tornando a detecção ainda mais robusta.

4.2 UNIFICAÇÃO ENTRE MACHINE LEARNING E DEEP LEARNING

De acordo com Inacio Junior e Cunha (2022), a aplicação da Inteligência Artificial contra tentativas de *phishing* permite que o sistema compreenda os padrões existentes de maneira mais aprofundada, resultando em classificações mais assertivas. Essa sinergia entre o aprendizado de máquina — focado em padrões estatísticos — e o Deep Learning — focado na semântica — possibilita que a tecnologia identifique golpes robustos que poderiam passar despercebidos por uma análise humana convencional.

Contudo, a literatura ressalta que a IA possui um caráter ambivalente da mesma forma que provê mecanismos de defesa, pode ser utilizada para a automação de golpes. De acordo com Morais e Branco (2023), embora a IA apresenta aplicações benéficas, seu potencial de uso indevido gera controvérsias. Nesse contexto, a tecnologia possibilita o desenvolvimento de mensagens customizadas em larga escala. Segundo Alves (2024), a Engenharia Social utiliza gatilhos emocionais, prática que se torna sofisticada quando o PLN é empregado para implementar o Spear *Phishing* de forma personalizada. Conforme Araújo (2024), essa integração permite uma análise semântica que conecta informações anteriores, dificultando o questionamento da vítima. Assim, observa-se que o cenário da segurança vive em constante adaptação, com métodos de prevenção e estratégias de ataque evoluindo simultaneamente.

5 CONSIDERAÇÕES FINAIS

As considerações finais do presente estudo permitem compreender que a Inteligência Artificial consolidou-se como uma ferramenta indispensável no combate ao *Phishing*, atuando de forma proativa na identificação e mitigação de ameaças digitais. A utilização de algoritmos como o XGBoost para o ajuste de erros e redes neurais do tipo LSTM para a análise profunda de textos possibilita que o sistema compreenda padrões complexos, classifique-os com precisão e aprenda a identificar novas táticas criminosas de maneira autônoma. Um ponto de destaque na análise é a dualidade inerente à tecnologia, evidenciando que a IA, embora fundamental para a defesa, também é explorada por agentes maliciosos para a sofisticação de ataques, o que torna imperativo que as ferramentas defensivas sejam capazes de detectar a atuação de outras inteligências artificiais em ambientes de rede. Dessa forma, a área da segurança cibernética configura-se em ciclos evolutivos constantes, buscando prever estratégias e garantir a proteção de usuários com diferentes níveis de letramento tecnológico.

A principal contribuição deste trabalho reside na demonstração de que a unificação entre técnicas de Machine Learning e as camadas de abstração do Deep Learning fundamenta uma abordagem defensiva mista e robusta. Essa sinergia permite o exame minucioso de URLs e e-mails suspeitos sob perspectivas estatísticas

e linguísticas, alertando o usuário sobre a presença de gatilhos de Engenharia Social, como o senso de urgência e o uso de parâmetros inseguros em links.

No entanto, identificou-se como lacuna de pesquisa a necessidade de maior investigação sobre a resposta desses modelos em tempo real diante de ataques gerados por IAs generativas. Especificamente, o surgimento de LLMs (Large Language Models) impõe um novo desafio, pois essas ferramentas permitem a criação de e-mails de *phishing* com alto grau de naturalidade e contexto, tornando-os virtualmente indistinguíveis de comunicações legítimas e desafiando a capacidade de detecção até mesmo dos filtros de PLN mais avançados. Como direcionamento futuro, será realizada a criação prática de algoritmos para validar os conceitos estabelecidos neste artigo, buscando testar a eficácia da integração entre aprendizado de máquina e redes neurais no enfrentamento dessas ameaças simuladas e altamente sofisticadas.

REFERÊNCIAS BIBLIOGRÁFICAS

A., G. *et al.* **Inteligência Artificial Aplicada à Cibersegurança: Soluções Estratégicas para um Ambiente Digital Resiliente**. Disponível em: https://immes.edu.br/wp-content/uploads/2025/04/Artigo_MATIZ_2023_Ciberseguranca.pdf?utm_source=chatgpt.com. Acesso em: 22 maio. 2026.

ALVES, Leonardo de Moura. **ENGENHARIA SOCIAL: ESTUDO DE ATAQUES E MÉTODOS DE PREVENÇÃO.** Orientador: Profa. Dra. Solange da Silva. 2024. TCC (Graduação) - Curso de Ciência da Computação, Escola Politécnica e de Artes, da Pontifícia Universidade Católica de Goiás, Goiás, 2024. Disponível em: https://repositorio.pucgoias.edu.br/jspui/handle/123456789/7976. acesso em: 11 mai. 2025.

ARAÚJO, David Pereira de. **Abordagens de Processamento de Linguagem Natural e Aprendizado Profundo para Classificação de Atos Administrativos de Diário Oficial.** Orientador: Henrique Coelho Fernandes. 2024. Dissertação (Mestrado) - Curso de Ciência da Computação, Universidade Federal de Uberlândia , Uberlândia, 2024. Disponível em: https://repositorio.ufu.br/bitstream/123456789/44590/1/AbordagensProcessamentoLinguagem.pdf . Acesso em: 25 nov. 2025.

BEZERRA, Paloma de Sousa; MILLIAN, Lucca Eiki Amarante; SILVA, Rodrigo Cardoso. **Estudo de Técnicas de *Phishing*: Métodos de Ataque e Estratégias de Defesa.** Faculdade de Computação e Informática (FCI), Universidade Presbiteriana Mackenzie, São Paulo, 2024. Disponível em: https://adelpha-

api.mackenzie.br/server/api/core/bitstreams/ea3bb8ff-17a5-4439-b98b-43725da49698/content. Acesso em: 4 dez. 2025.

BUTT, Umer Ahmed *et al. Cloud-based email phishing attack using machine and deep learning algorithm*. ***Complex & Intelligent Systems***, v. 9, n. 3, p. 3043–3070, 2023.

COUTINHO, VINICIUS MACHADO. **DETECÇÃO DE PÁGINAS DE *PHISHING* UTILIZANDO APRENDIZADO DE MÁQUINA.** Orientador: Kelton Augusto Pontara da Costa. 2023. TCC (Graduação) - Curso de Ciência da Computação, Universidade Estadual Paulista "Júlio de Mesquita Filho", Bauru, 2023. Disponível em: https://repositorio.unesp.br/server/api/core/bitstreams/6818f06d-40ea-4d62-b967-d815e46bf056/content. acesso em: 26 nov. 2025.

DE SÁ DOUGLAS VIEIRA BARBOZA, Rodrigo Figueiredo Costa de Oliveira Sidney Loyola. **Simulação de *phishing* como ferramenta de conscientização em cibersegurança: um estudo de caso em ambiente corporativo**. Disponível em: <https://ojs.revistacontribuciones.com/ojs/index.php/clcs/article/view/19198>. Acesso em: 22 maio. 2026.

GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. **Métodos de Pesquisa.** Rio Grande do Sul. 2009. Disponível em: https://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf. Acesso em: 11 mai. 2025.

HOSAKI, Gabriel Yuri; RIBEIRO, Douglas Francisco. DEEP LEARNING. **Revista de Gestão e Estratégia**, Assis, ano 2021, Disponível em: https://ric.cps.sp.gov.br/bitstream/123456789/5060/1/DEEP-LEARNING.pdf. Acesso em: 11 mai. 2025.

INACIO JUNIOR, Marcos Antonio; CUNHA, Alberane Lucio Thiago da. COMO A INTELIGÊNCIA ARTIFICIAL PODE SER APLICADA CONTRA AS TENTATIVAS DE *PHISHING*. **COMO A INTELIGÊNCIA ARTIFICIAL PODE SER APLICADA CONTRAAS TENTATIVAS DE PHISHING**, [s. l.], 30 nov. 2022 Disponível em: http://repositorio.unis.edu.br/handle/prefix/2483. Acesso em: 11 mai. 2025.

JUDE OSAMOR, MOSES APROFIN ASHAWA, JACKIE RILEY, PIUS OWOH. ***Real-Time Detection of Phishing Emails Using XG Boost Machine Learning Technique***. Disponível em: <https://www.researchgate.net/publication/393647677_Real-Time_Detection_of_Phishing_Emails_Using_XG_Boost_Machine_Learning_Technique>. Acesso em: 23 maio. 2026.

MORAIS, Prof. Me. Flávio Daniel Borges de; BRANCO, Prof. Me. Valdec Romero Castelo. A Inteligência Artificial: conceitos, aplicações e controvérsias. **XX SICI**, Guarujá, ano 2023, Disponível em: https://www.unaerp.br/pesquisa/anais-de-congressos-unaerp/sici. Acesso em: 11 mai. 2025.

SAMPAIO, Tuane Bazanella. **Metodologia da pesquisa.** Santa Maria, Rio Grande do Sul. 2022. Disponível em:

https://repositorio.ufsm.br/bitstream/handle/1/26138/MD_Metodologia_da_Pesquisa.pdf?sequence=1&isAllowed=y. Acesso em: 11 mai. 2025.