REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

DESENVOLVIMENTO DE SISTEMAS INTELIGENTES PARA DETECÇÃO DE VULNERABILIDADES EM TEMPO REAL COM APLICAÇÕES EM CIBERSEGURANÇA

DEVELOPMENT OF INTELLIGENT SYSTEMS FOR REAL-TIME VULNERABILITY DETECTION WITH APPLICATIONS IN CYBERSECURITY

Ciências Exatas e da Terra, Engenharias • 06/04/2026

REGISTRO DOI: 10.70773/revistatopicos/775413946

Rômulo Ferreira dos Santos[1]

Paulo Cesar Rodrigues Borges[2]

Douglas Alves Soares[3]

Marcus Dhilermando Hora de Souza[4]

Cássio Natan Santos Ferreira [5]

RESUMO

A crescente complexidade dos sistemas digitais, associada à ampliação da superfície de ataque em redes corporativas, aplicações em nuvem, dispositivos conectados e cadeias de suprimento de software, tornou insuficiente a dependência exclusiva de mecanismos tradicionais de detecção de vulnerabilidades. Nesse contexto, a integração entre cibersegurança, engenharia de software e sistemas inteligentes emerge como alternativa estratégica para ampliar a capacidade de monitoramento, priorização e resposta diante de fraquezas exploráveis em tempo oportuno. Este artigo analisa o desenvolvimento de sistemas inteligentes voltados à detecção de vulnerabilidades em tempo real, com ênfase em aplicações corporativas, articulando aprendizado de máquina, análise estática e dinâmica, inteligência sobre ameaças, classificação de fraquezas de software e monitoramento contínuo de eventos de segurança. A partir de revisão sistemática qualitativa e análise documental, o estudo discute os fundamentos técnicos desse campo, distingue vulnerabilidade, fraqueza e incidente, e propõe uma arquitetura híbrida para detecção e mitigação proativa em ambientes organizacionais. A literatura mostra que modelos inteligentes podem ampliar a capacidade de identificação de padrões complexos, reduzir tempo de triagem e apoiar a priorização de correções, especialmente quando combinados a estruturas como CWE, CVSS, ATT&CK, catálogos de vulnerabilidades exploradas e práticas de desenvolvimento seguro. Contudo, persistem desafios importantes, como dados desbalanceados, alta taxa de falso positivo, dificuldade de generalização entre bases, opacidade dos modelos, deriva de desempenho e exposição a ataques adversariais contra os próprios mecanismos de inteligência. Conclui-se que a detecção em tempo real exige abordagem multicamadas e governança contínua, em que a inteligência artificial atue como mecanismo

complementar e auditável, e não como substituta isolada dos controles clássicos de segurança, das práticas seguras de desenvolvimento e da resposta coordenada a incidentes.
Palavras-chave: engenharia de software; cibersegurança; inteligência artificial; aprendizado de máquina; vulnerabilidades digitais; detecção em tempo real.

ABSTRACT

The increasing complexity of digital systems, combined with the expansion of the attack surface across corporate networks, cloud applications, connected devices, and software supply chains, has made exclusive reliance on traditional vulnerability detection mechanisms insufficient. In this context, the integration of cybersecurity, software engineering, and intelligent systems emerges as a strategic alternative to enhance monitoring, prioritization, and response capabilities for exploitable weaknesses in a timely manner. This article analyzes the development of intelligent systems for real-time vulnerability detection, with emphasis on corporate applications, by combining machine learning, static and dynamic analysis, threat intelligence, software weakness classification, and continuous monitoring of security events. Based on a qualitative systematic review and documentary analysis, the study discusses the technical foundations of this field, distinguishes vulnerability, weakness, and incident, and proposes a hybrid architecture for proactive detection and mitigation in organizational environments. The literature shows that intelligent models can expand the capacity to identify complex patterns, reduce triage time, and support remediation prioritization, especially when combined with structures such as CWE, CVSS, ATT&CK, catalogs of known exploited vulnerabilities, and secure software development practices. However, important challenges remain, including

imbalanced datasets, high false-positive rates, difficulty generalizing across datasets, model opacity, performance drift, and exposure to adversarial attacks against the intelligence mechanisms themselves. It is concluded that real-time detection requires a multilayered approach and continuous governance, in which artificial intelligence acts as a complementary and auditable mechanism rather than an isolated substitute for classical security controls, secure development practices, and coordinated incident response.
Keywords: software engineering; cybersecurity; artificial intelligence; machine learning; digital vulnerabilities; real-time detection.

1. INTRODUÇÃO

A sociedade contemporânea atravessa um processo de transformação digital sem precedentes, caracterizado pela intensificação do uso de sistemas computacionais em praticamente todos os setores da vida econômica, institucional e social. Organizações públicas e privadas passaram a depender, de modo crescente, de redes corporativas complexas, aplicações em nuvem, dispositivos móveis, plataformas integradas, sistemas ciberfísicos, internet das coisas, bancos de dados distribuídos e cadeias de suprimento digitais altamente interconectadas. Essa expansão da infraestrutura tecnológica produziu ganhos significativos em produtividade, conectividade, escalabilidade e automação, mas também ampliou de forma expressiva a superfície de ataque disponível para agentes maliciosos. Em consequência, a cibersegurança deixou de ser uma preocupação periférica, restrita a setores especializados da tecnologia da informação, e passou a constituir tema central para a continuidade operacional, a proteção de ativos críticos, a segurança de dados sensíveis, a conformidade

regulatória e a própria confiança no funcionamento das organizações digitais.

Nesse contexto, a vulnerabilidade digital assume papel central na análise dos riscos cibernéticos. Vulnerabilidades podem estar associadas a falhas de codificação, erros de configuração, fraquezas arquiteturais, dependências inseguras, mecanismos de autenticação inadequados, permissões excessivas, interfaces mal protegidas, atualizações não aplicadas, integração deficiente entre sistemas e diversas outras condições que criam oportunidades de exploração. Diferentemente de uma visão simplificada que associa o problema apenas a “erros de software”, a literatura e os referenciais técnicos mostram que a vulnerabilidade é resultado de uma interação complexa entre tecnologia, processos, design, contexto operacional e capacidade adversária. Uma mesma fraqueza pode ter impacto reduzido em determinado ambiente e representar risco extremo em outro, a depender da criticidade do ativo afetado, da exposição à internet, da presença de controles compensatórios e da existência de exploração conhecida.

A crescente sofisticação das ameaças cibernéticas agrava ainda mais esse cenário. Ataques contemporâneos já não se limitam à exploração manual de falhas isoladas, mas frequentemente envolvem automação, reconhecimento em larga escala, exploração coordenada de múltiplas vulnerabilidades, uso de inteligência sobre alvos, abuso de credenciais, ataques à cadeia de suprimento e técnicas de evasão destinadas a contornar mecanismos tradicionais de defesa. Além disso, a velocidade com que novas vulnerabilidades são divulgadas, catalogadas e exploradas tem imposto desafios significativos às equipes de segurança. Em muitas organizações, o volume de alertas, a diversidade de ativos, a pressão por

disponibilidade contínua dos serviços e a escassez de profissionais especializados tornam inviável uma abordagem exclusivamente manual ou reativa. Assim, a gestão de vulnerabilidades passa a exigir novos paradigmas, capazes de operar em escala, com agilidade e capacidade analítica compatível com a complexidade do ambiente digital atual.

Historicamente, a identificação de vulnerabilidades esteve fortemente apoiada em mecanismos clássicos, como análise estática de código, análise dinâmica, varreduras de configuração, testes de penetração, revisão manual e sistemas baseados em assinaturas. Tais abordagens continuam sendo fundamentais e não podem ser abandonadas, pois constituem a base consolidada da prática de segurança ofensiva e defensiva. No entanto, sua atuação isolada apresenta limitações cada vez mais evidentes. Ferramentas de análise estática, por exemplo, podem gerar grande quantidade de alertas, nem sempre contextualizados em relação ao risco real. Testes dinâmicos e varreduras dependem de janela operacional, escopo de cobertura e frequência de execução. Testes de penetração, embora extremamente valiosos, têm custo elevado, dependem de expertise especializada e são, em muitos casos, episódicos. Sistemas baseados unicamente em padrões conhecidos tendem a apresentar menor capacidade de generalização diante de novas variantes de ataque, novos artefatos maliciosos ou formas inéditas de exploração. Em síntese, os métodos tradicionais seguem necessários, mas já não são suficientes, quando empregados de modo isolado, para responder à velocidade e à escala das ameaças digitais.

É nesse ponto que os sistemas inteligentes ganham relevância estratégica. O avanço recente da inteligência artificial, do

aprendizado de máquina e do aprendizado profundo abriu novas possibilidades para a análise automatizada de grandes volumes de dados, identificação de padrões complexos, correlação entre eventos heterogêneos, reconhecimento de anomalias e apoio à decisão em contextos de elevada incerteza. Na cibersegurança, essas capacidades podem ser aplicadas à análise de código-fonte, à classificação de vulnerabilidades, à detecção de tráfego suspeito, ao reconhecimento de comportamentos anômalos em redes corporativas, ao direcionamento de fuzzing, à priorização de correções e ao monitoramento contínuo de ambientes operacionais. A hipótese central que sustenta esse movimento é que algoritmos inteligentes são capazes de perceber relações e sinais que escapam, parcial ou totalmente, aos métodos tradicionais baseados em regras fixas ou observação manual. Isso não significa, contudo, que a inteligência artificial forneça solução mágica ou substitua automaticamente a expertise humana. Ao contrário, seu valor depende da forma como é integrada aos processos, aos controles existentes e à governança organizacional.

A adoção de inteligência artificial em cibersegurança se insere, portanto, em uma agenda mais ampla de construção de capacidades preditivas e adaptativas. Em vez de esperar pela ocorrência de incidentes já consumados, as organizações buscam antecipar fragilidades, estimar risco de exploração, classificar prioridades e reduzir o tempo de resposta. O conceito de detecção em tempo real emerge justamente desse esforço de redução de latência entre exposição, percepção do risco e ação corretiva. Entretanto, é necessário esclarecer que o termo “tempo real” não deve ser interpretado de forma ingênua ou absoluta. Em cibersegurança, detectar em tempo real não significa descobrir instantaneamente qualquer vulnerabilidade no momento exato em

que ela surge. Significa construir uma capacidade contínua de monitoramento, correlação e análise que opere com rapidez suficiente para reduzir o intervalo entre o aparecimento de sinais críticos e a tomada de decisão defensiva. Trata-se, portanto, de uma propriedade sistêmica da organização e de sua arquitetura de segurança, e não apenas de uma funcionalidade isolada de determinada ferramenta.

A discussão sobre vulnerabilidades em tempo real exige também uma distinção conceitual rigorosa entre fraqueza, vulnerabilidade, exploração e incidente. No campo técnico, uma fraqueza corresponde a uma deficiência de projeto, implementação, configuração ou arquitetura que pode favorecer o surgimento de vulnerabilidades. A vulnerabilidade, por sua vez, representa uma condição explorável que pode comprometer confidencialidade, integridade, disponibilidade ou outros objetivos de segurança. A exploração ocorre quando um agente adversário utiliza essa condição para obter acesso indevido, executar código, elevar privilégios, exfiltrar dados, comprometer serviços ou produzir dano operacional. Já o incidente corresponde à materialização de um evento de segurança com consequências observáveis e relevantes para a organização. Essa distinção é importante porque sistemas inteligentes podem atuar em diferentes momentos dessa cadeia: podem identificar fraquezas antes da implantação, vulnerabilidades conhecidas em componentes já instalados, sinais de tentativa de exploração em rede ou comportamentos anômalos que indiquem comprometimento em curso.

Em ambientes corporativos, essa visão encadeada é particularmente relevante porque a segurança depende da convergência entre desenvolvimento seguro, gestão de ativos, monitoramento

operacional e resposta coordenada. Uma vulnerabilidade em uma biblioteca de terceiros, por exemplo, pode não representar risco imediato se o componente não estiver presente no inventário real da organização, se a funcionalidade vulnerável não estiver exposta ou se existirem controles compensatórios eficazes. Em contrapartida, uma vulnerabilidade classificada inicialmente como moderada pode se tornar crítica caso seja amplamente explorada por agentes de ameaça, esteja presente em um ativo essencial e permita movimentação lateral ou exfiltração de dados estratégicos. Isso demonstra que a detecção eficaz não depende apenas de saber “o que existe”, mas de compreender “onde existe”, “como pode ser explorado”, “por quem pode ser explorado” e “qual impacto pode produzir no contexto específico do negócio”. Sistemas inteligentes podem contribuir justamente nesse nível de contextualização, desde que alimentados por dados adequados e integrados a estruturas sólidas de priorização de risco.

A literatura recente sobre aprendizado de máquina aplicado à detecção de vulnerabilidades mostra um campo em rápida expansão. Estudos têm investigado o uso de representações sintáticas e semânticas de código, árvores de sintaxe abstrata, grafos de fluxo de controle, embeddings de programas, redes neurais profundas, classificadores supervisionados e técnicas híbridas para identificar trechos potencialmente vulneráveis antes da exploração em produção. Paralelamente, a área de detecção de intrusão baseada em inteligência artificial tem avançado na análise de tráfego, registros de sistema, comportamento de usuários e anomalias operacionais. Há ainda o crescimento do fuzzing orientado por aprendizado de máquina, que busca maximizar cobertura de execução e descobrir falhas inéditas de forma mais eficiente. Esses avanços demonstram que a inteligência artificial

possui potencial real para contribuir com a cibersegurança, sobretudo quando aplicada à ampliação da visibilidade e da capacidade analítica em escala.

Apesar desse potencial, a adoção de métodos inteligentes também impõe desafios substanciais. Um dos mais relevantes diz respeito à qualidade e à representatividade dos dados utilizados para treinamento e validação dos modelos. Bases de vulnerabilidades, repositórios de código e conjuntos de tráfego frequentemente apresentam desbalanceamento, rotulagem imperfeita, duplicidades, ruído e viés de seleção. Isso pode produzir modelos aparentemente eficazes em ambiente experimental, mas frágeis quando submetidos a cenários reais e heterogêneos. Outro problema recorrente é a interpretabilidade. Em cibersegurança, alertas precisam ser compreendidos por analistas, desenvolvedores, gestores e auditores; portanto, decisões automatizadas opacas reduzem confiança e dificultam validação operacional. Soma-se a isso o risco de falsos positivos em escala, que pode gerar fadiga analítica e comprometer a capacidade de resposta das equipes. Assim, o desenvolvimento de sistemas inteligentes eficazes não depende apenas de obter boa acurácia em laboratório, mas de garantir robustez, transparência, governança e valor operacional em condições reais de uso.

Há ainda uma camada adicional de complexidade: os próprios sistemas inteligentes se tornam alvos potenciais de ataque. Modelos de aprendizado de máquina podem sofrer manipulação adversarial, envenenamento de dados, evasão por amostras maliciosamente elaboradas, roubo de modelo e comprometimento da cadeia de suprimento. Em outras palavras, o mecanismo projetado para fortalecer a defesa pode, ele próprio, introduzir novas superfícies de

risco. Isso exige que a inteligência artificial aplicada à cibersegurança seja concebida de forma segura desde a origem, com documentação adequada, controle de versões, monitoramento de desempenho, auditoria contínua, validação de insumos e avaliação de robustez frente a comportamentos adversariais. A segurança da inteligência torna-se, portanto, parte inseparável da inteligência para segurança.

No plano organizacional, o uso de sistemas inteligentes também precisa ser compreendido como questão de governança. A simples aquisição ou implementação de algoritmos avançados não garante melhoria efetiva da postura de segurança. É necessário definir fluxos de decisão, critérios de confiança, responsabilidades sobre os modelos, integração com centros de operações de segurança, interfaces com equipes de desenvolvimento e mecanismos claros de retroalimentação dos achados para o ciclo de vida do software. Do contrário, a organização corre o risco de acumular ferramentas sem integração, modelos sem manutenção, alertas sem tratamento e iniciativas analíticas desconectadas das prioridades reais do negócio. Assim, a maturidade em cibersegurança inteligente envolve não apenas capacidade técnica, mas também capacidade institucional.

Diante desse panorama, torna-se evidente a relevância científica e prática de investigar como desenvolver sistemas inteligentes voltados à detecção de vulnerabilidades em tempo real com aplicações em cibersegurança. O tema é particularmente importante porque se localiza na interseção entre engenharia de software, inteligência artificial, ciência de dados, arquitetura de sistemas, gestão de risco e segurança operacional. Ao mesmo tempo em que demanda sofisticação técnica, exige sensibilidade para compreender limitações reais, riscos de implementação e

condicionantes organizacionais. Não basta perguntar se um algoritmo pode detectar uma vulnerabilidade; é preciso perguntar em que contexto, com quais dados, com qual grau de explicação, com que custo de operação, com que impacto sobre falsos positivos, com que aderência ao ambiente corporativo e com que integração aos processos já existentes.

É justamente essa necessidade de abordagem integrada que orienta o presente estudo. O trabalho parte do reconhecimento de que a cibersegurança moderna não pode mais depender exclusivamente de mecanismos pontuais ou reativos, e de que a inteligência artificial, embora não seja solução autossuficiente, oferece recursos promissores para o fortalecimento da detecção proativa de vulnerabilidades. Ao discutir fundamentos conceituais, possibilidades técnicas, desafios metodológicos e implicações organizacionais, esta pesquisa busca contribuir para uma compreensão mais crítica, sólida e operacional do tema. A proposta não é apenas descrever tecnologias emergentes, mas analisar como elas podem ser incorporadas de forma responsável e efetiva à proteção de redes corporativas e sistemas digitais complexos.

Dessa forma, o objetivo deste artigo é analisar o desenvolvimento de sistemas inteligentes para detecção de vulnerabilidades em tempo real, identificando seus fundamentos técnicos, suas potencialidades para a proteção proativa de dados e infraestruturas, seus principais desafios de implementação e seus limites em termos de governança, confiabilidade e integração organizacional. Parte-se da premissa de que a detecção inteligente de vulnerabilidades deve ser concebida como capacidade sistêmica e contínua, sustentada por múltiplas fontes de dados, técnicas analíticas complementares, estruturas reconhecidas de classificação e priorização de risco e

participação ativa de especialistas humanos no processo de validação e resposta. Ao final, busca-se demonstrar que o futuro da cibersegurança corporativa depende menos da adoção isolada de algoritmos sofisticados e mais da articulação entre inteligência computacional, engenharia segura, observabilidade operacional e responsabilidade institucional.

2. METODOLOGIA

O estudo foi desenvolvido como **revisão sistemática qualitativa**, complementada por **análise documental técnica**, com foco em dois eixos: o primeiro, voltado à literatura acadêmica sobre detecção automática de vulnerabilidades com aprendizado de máquina; o segundo, centrado em referenciais institucionais e normativos relevantes para vulnerabilidade, risco cibernético, desenvolvimento seguro, classificação de fraquezas e governança do ciclo de vida dos modelos.

No eixo acadêmico, foram priorizados estudos de síntese e trabalhos de referência sobre detecção de vulnerabilidades em código-fonte, representação semântica de programas, aprendizado profundo para identificação de falhas, fuzzing orientado por aprendizado de máquina e avaliação empírica de modelos. Entre os trabalhos selecionados, destacam-se a revisão sistemática de **Harzevili et al. (2024)**, o estudo de **Marjanov, Pashchenko e Massacci (2022)** sobre o que funciona e o que ainda não funciona em aprendizado de máquina aplicado à detecção de vulnerabilidades em código, o estudo empírico de **Semasaba et al. (2022/2023)** sobre representações de código e desempenho de modelos e o trabalho de **Bilgin et al. (2020)** sobre predição de vulnerabilidades a partir do código-fonte. Também foi considerada a revisão sistemática de

Chafjiri et al. (2024) sobre fuzzing baseado em aprendizado de máquina.

No eixo técnico-institucional, foram selecionados documentos do **NIST**, **CISA**, **MITRE**, **FIRST** e **OWASP**, por sua relevância prática e normativa. Foram usados o **CSF 2.0** para gestão de risco e monitoramento contínuo, o **SSDF** para desenvolvimento seguro, os materiais do **SAMATE/SATE** sobre avaliação de ferramentas de análise estática, o **CWE** para classificação de fraquezas, o **ATT&CK** para contextualização de técnicas adversárias, o **CVSS v4.0** para severidade e priorização, o catálogo de **vulnerabilidades conhecidamente exploradas** da **CISA** e as taxonomias de riscos para sistemas de aprendizado de máquina e aprendizado de máquina adversarial.

A análise foi conduzida por síntese narrativa crítica. Em vez de comparar apenas desempenho numérico de modelos em bases experimentais, o estudo procurou responder a quatro perguntas: **(i)** o que significa detectar vulnerabilidades "em tempo real" em contexto corporativo; **(ii)** quais fontes de dados e métodos devem ser integrados; **(iii)** quais limitações impedem implantação segura em produção; e **(iv)** como estruturar arquitetura prática de uso organizacional. O resultado dessa síntese foi a formulação de um modelo conceitual híbrido para detecção contínua e resposta orientada a risco.

3. RESULTADOS E DISCUSSÃO

3.1. A Detecção em Tempo Real Como Problema Multicamada

A principal constatação da revisão é que "detecção de vulnerabilidades em tempo real" não deve ser compreendida de

forma restrita como descoberta instantânea de defeitos inéditos no código, mas como **capacidade operacional contínua** de identificar fraquezas, correlacionar sinais de exploração e priorizar correções com baixa latência. Em termos práticos, isso envolve integrar análise de código, composição de software, telemetria de rede, registros de eventos, comportamento de processos, inteligência sobre ameaças e catálogos de exploração ativa. A literatura e os referenciais técnicos convergem para a ideia de que a defesa eficiente depende de monitoramento contínuo, análise contextual e resposta coordenada.

Essa visão multicamada resolve uma ambiguidade frequente na literatura aplicada. Muitas pesquisas tratam “detecção de vulnerabilidades” como classificação de trechos de código vulneráveis. Outras tratam “detecção em tempo real” como identificação de intrusões ou anomalias em tráfego e eventos. Em ambientes corporativos reais, os dois domínios se encontram: uma fraqueza de software catalogável por **CWE** ou **CVE** pode ser explorada por técnicas mapeáveis em **ATT&CK**, e o valor da inteligência reside justamente em conectar a fraqueza potencial à probabilidade e ao impacto de exploração efetiva. Por isso, sistemas inteligentes mais úteis são aqueles que operam na transição entre engenharia de software, gestão de ativos e segurança operacional.

3.2. Métodos Inteligentes Mais Promissores

A revisão acadêmica recente indica quatro grandes famílias de métodos. A primeira é a **análise de código baseada em aprendizado de máquina**, que utiliza tokens, árvores de sintaxe abstrata, grafos de fluxo de controle, grafos de fluxo de dados e embeddings semânticos para classificar funções, arquivos ou trechos de código como vulneráveis ou não vulneráveis. Harzevili et

al. mostram que esse campo cresceu rapidamente, enquanto Marjanov et al. destacam que a principal promessa dessas abordagens é capturar padrões semânticos mais profundos do que regras tradicionais isoladas. Bilgin et al. também demonstram a viabilidade de extrair representações estruturadas do código para predição antes da liberação do software.

A segunda família é a **detecção de intrusão e anomalias em tráfego e eventos operacionais**, em que modelos supervisionados e não supervisionados aprendem padrões de normalidade e distinguem comportamentos anômalos em rede, hosts ou serviços. Embora esse campo seja tradicionalmente associado à detecção de ataques e não de vulnerabilidades em si, sua relevância para o tema é alta, pois fornece sinais indiretos de exploração, movimentação lateral, varredura e abuso de configurações frágeis. O NIST, ao definir sistemas de detecção e prevenção de intrusão, reforça essa lógica de monitoramento contínuo de eventos; estudos recentes sobre estruturas de aprendizado de máquina em tempo real também apontam que a eficácia prática depende de integração com telemetria de produção.

A terceira família é o **fuzzing orientado por aprendizado de máquina**, que busca guiar geração de entradas de teste e explorar caminhos de execução de forma mais eficiente do que técnicas puramente aleatórias. A revisão sistemática de Chafjiri et al. identifica esse campo como uma das frentes mais promissoras para encontrar comportamentos inesperados e falhas ainda não formalmente catalogadas, sobretudo quando o objetivo é ampliar cobertura de execução e acelerar descoberta de falhas difíceis de reproduzir. Essa linha é particularmente relevante quando se discute

detecção proativa de vulnerabilidades antes da exploração em produção.

A quarta família é a **priorização inteligente de vulnerabilidades**, que combina severidade, contexto de ativo, exposição, exploração observada e criticidade de negócio. Aqui, o uso conjunto de **CVSS v4.0**, catálogo **KEV** da **CISA**, inventários de componentes e inteligência contextual permite reduzir o problema clássico da "fila infinita" de remediação. Em ambientes corporativos, detectar é insuficiente; é necessário ordenar o que corrigir primeiro. A CISA recomenda explicitamente revisão e monitoramento do catálogo de vulnerabilidades conhecidamente exploradas, enquanto o CVSS v4.0 acrescenta dimensões de ameaça e ambiente à análise de severidade.

3.3. Limitações Técnicas dos Modelos

Apesar do avanço, a literatura é clara ao afirmar que o desempenho de sistemas inteligentes ainda enfrenta barreiras importantes. A primeira delas é a **qualidade das bases de dados**. Muitas bases usadas em pesquisa são construídas com rótulos derivados de ferramentas automáticas, com forte desbalanceamento entre classes, duplicidade de amostras ou distância considerável entre o ambiente experimental e o código corporativo real. Harzevili et al. identificam esse problema como recorrente, e Semasaba et al. mostram que o desempenho dos modelos varia significativamente conforme a representação do código, a estratégia de aprendizado e o tratamento de desbalanceamento.

A segunda limitação é a **taxa de falso positivo**. Em segurança corporativa, alertas excessivos podem inviabilizar a adoção do

sistema, gerar fadiga analítica e retardar respostas efetivamente críticas. O histórico do programa **SATE**, do NIST, evidencia que a avaliação de ferramentas de análise estática sempre esteve ligada não apenas à descoberta de fraquezas, mas também à utilidade prática dos relatórios produzidos. Em outras palavras, um sistema inteligente que detecta muito, mas explica pouco e prioriza mal, tende a ser operacionalmente fraco.

A terceira limitação é a **generalização**. Marjanov et al. alertam que ainda há lacunas importantes para que o aprendizado de máquina em detecção de vulnerabilidades seja considerado maduro: representações semânticas complexas, robustez fora da base de treino, interpretabilidade e comparação justa entre métodos permanecem desafios centrais. O estudo empírico de Steenhoek et al. também reforça que, embora modelos profundos possam superar ferramentas tradicionais em alguns conjuntos de dados, ainda falta compreensão suficiente sobre robustez, depuração e implantação confiável em cenários reais.

A quarta limitação envolve a **explicabilidade e a auditabilidade**. Em contexto corporativo, um alerta de vulnerabilidade precisa ser compreensível para desenvolvedores, equipes de segurança e gestores de risco. A proposta de **Model Cards**, de Mitchell et al., embora originalmente mais ampla, é útil aqui porque oferece lógica de documentação do propósito, limitações, desempenho e escopo de modelos. Sem isso, o uso de inteligência artificial em cibersegurança corre o risco de produzir decisões pouco transparentes e difíceis de contestar ou melhorar.

3.4. Riscos Específicos dos Próprios Sistemas Inteligentes

Um ponto decisivo da literatura recente é que sistemas inteligentes de segurança também precisam ser protegidos. O relatório do **NIST** sobre aprendizado de máquina adversarial organiza ataques por fase do ciclo de vida e por objetivo do atacante, enquanto a **OWASP** destaca riscos como manipulação de entrada, envenenamento de dados, inversão de modelo, roubo de modelo e ataques à cadeia de suprimento de inteligência artificial. Isso significa que um detector mal governado pode ser enganado, degradado ou manipulado, exatamente no momento em que deveria proteger a organização.

Por isso, a implantação de sistemas inteligentes em cibersegurança exige monitoramento contínuo de desempenho, validação periódica, rastreabilidade de dados, controle de versões, supervisão humana e critérios para desativação ou reentrenamento do modelo. O **AI RMF** do NIST e seus materiais complementares insistem em governança, mensuração e gestão contínua dos riscos de IA, o que é particularmente importante quando o modelo influencia triagem, priorização ou resposta a incidentes.

3.5. Proposta de Arquitetura Híbrida para Uso Corporativo

Com base na síntese realizada, propõe-se uma **arquitetura híbrida de detecção em tempo real** composta por seis camadas integradas.

A **primeira camada** é a de **coleta contínua**, reunindo código-fonte, resultados de análise estática, composição de software, inventário de ativos, registros de eventos, tráfego de rede, telemetria de endpoint, dados de configuração e inteligência externa sobre ameaças. O uso de **SBOM** é especialmente relevante nessa camada porque aumenta a visibilidade sobre componentes e dependências, acelerando a

identificação e a remediação de vulnerabilidades na cadeia de software.

A **segunda camada** é a de **normalização e enriquecimento**, em que os achados são associados a **CWE**, **CVE**, **CVSS v4.0**, **ATT&CK** e catálogo **KEV**. Esse enriquecimento permite passar de um alerta bruto para uma avaliação contextual de severidade, explorabilidade e impacto potencial. Em vez de uma lista genérica de falhas, a organização passa a ter evidências estruturadas para priorizar correções e contenções.

A **terceira camada** é a de **motor analítico híbrido**, combinando regras determinísticas, modelos supervisionados para classificação de vulnerabilidades, modelos não supervisionados para anomalias, correlação temporal de eventos e fuzzing inteligente em ciclos de validação. A revisão mostra que o melhor desempenho operacional tende a surgir da combinação de métodos, e não da dependência exclusiva de um único modelo.

A **quarta camada** é a de **pontuação e decisão**, na qual cada achado recebe uma pontuação de risco composta por severidade, exploração conhecida, exposição do ativo, criticidade do serviço e confiança do modelo. Aqui, a lógica do **CVSS v4.0** e a priorização baseada em vulnerabilidades já exploradas observadas pela **CISA** tornam-se extremamente úteis.

A **quinta camada** é a de **resposta e mitigação**, integrando abertura de chamados, isolamento de ativos, correções de configuração, recomendação de atualização, bloqueios preventivos e alimentação das equipes de desenvolvimento seguro. Essa camada dialoga diretamente com o **CSF 2.0**, que organiza os resultados de

segurança em governança, identificação, proteção, detecção, resposta e recuperação.

A **sexta camada** é a de **governança do ciclo de vida**, responsável por documentação de modelos, monitoramento de deriva, auditoria, testes contra ataques adversariais, versionamento e alinhamento às práticas do **SSDF**. Essa última camada é decisiva porque transforma um experimento de inteligência artificial em um sistema corporativo confiável e sustentado por engenharia de software segura.

3.6. Implicações para Redes Corporativas

Nas redes corporativas, a contribuição mais valiosa dos sistemas inteligentes não está apenas em "achar mais falhas", mas em **reduzir o tempo entre exposição, detecção, priorização e resposta**. Isso é especialmente relevante quando se considera que vulnerabilidades conhecidas continuam sendo amplamente exploradas e que a priorização baseada em risco real tende a ser mais eficiente do que a correção cega por lista extensa. A CISA, ao recomendar monitoramento do catálogo KEV, reforça exatamente esse ponto: nem toda vulnerabilidade merece a mesma urgência operacional.

Além disso, a integração entre segurança de desenvolvimento e segurança operacional fortalece a resiliência organizacional. O **SSDF** recomenda incorporar práticas de desenvolvimento seguro ao ciclo de vida do software, e isso dialoga diretamente com a lógica aqui defendida: sistemas inteligentes devem alimentar tanto a operação quanto a engenharia, retroalimentando correções, revisões de código, melhoria arquitetural e redução de reincidência de fraquezas.

4. CONCLUSÃO

A análise desenvolvida ao longo deste estudo permite afirmar que a detecção de vulnerabilidades em tempo real, apoiada por sistemas inteligentes, representa uma das frentes mais promissoras e estratégicas da cibersegurança contemporânea. Em um cenário caracterizado pela intensificação da transformação digital, pela ampliação da superfície de ataque, pela crescente sofisticação dos agentes maliciosos e pela interdependência entre software, infraestrutura, serviços em nuvem, dispositivos conectados e cadeias de suprimento tecnológicas, a adoção de métodos convencionais isolados já não se mostra suficiente para garantir níveis satisfatórios de proteção. A cibersegurança deixou de ser uma atividade meramente reativa, centrada em correções posteriores a incidentes, e passou a exigir abordagens preventivas, contínuas, preditivas e adaptativas. É justamente nesse contexto que os sistemas inteligentes baseados em inteligência artificial e aprendizado de máquina se consolidam como instrumentos relevantes para fortalecer a capacidade das organizações de identificar, classificar, priorizar e mitigar vulnerabilidades antes que estas sejam exploradas com impacto significativo.

Uma das principais conclusões deste estudo é que o valor dos sistemas inteligentes em cibersegurança não reside apenas na automação da análise técnica, mas na sua capacidade de operar como mecanismo de ampliação cognitiva das equipes humanas e dos controles institucionais. Em outras palavras, tais sistemas não devem ser compreendidos como substitutos absolutos da análise especializada, da governança de risco ou da engenharia segura, mas como elementos de apoio à tomada de decisão em ambientes altamente complexos e dinâmicos. Quando integrados de forma

adequada a fluxos de telemetria, análise estática e dinâmica, inventário de ativos, inteligência sobre ameaças, classificação de fraquezas de software e práticas de resposta coordenada, esses sistemas aumentam significativamente a velocidade com que sinais relevantes são percebidos e transformados em ação defensiva. O ganho organizacional mais importante, portanto, não é apenas “detectar mais”, mas detectar melhor, com maior contexto, menor latência e maior capacidade de priorização.

Nesse sentido, ficou evidenciado que a noção de detecção em tempo real precisa ser compreendida em sentido técnico mais amplo e mais realista. Não se trata da ideia simplificada de descobrir instantaneamente qualquer falha de software, como se a tecnologia inteligente pudesse, sozinha, eliminar a incerteza inerente ao ambiente digital. Trata-se, antes, da construção de uma capacidade contínua de observação, correlação e reação, capaz de reduzir o intervalo entre o surgimento de sinais relevantes, a identificação de exposição potencial, a compreensão do risco efetivo e a implementação de medidas de contenção ou correção. Essa visão é crucial porque impede expectativas irreais sobre a inteligência artificial e reposiciona o debate no campo da engenharia de sistemas de segurança. O desafio central não é criar um detector infalível, mas desenhar ecossistemas defensivos em que múltiplas fontes de evidência sejam combinadas em uma arquitetura operacional robusta.

A literatura analisada também permite concluir que os métodos inteligentes mais eficazes são aqueles que trabalham em regime de complementaridade. A detecção de vulnerabilidades em código-fonte por aprendizado de máquina, a identificação de anomalias comportamentais em redes e endpoints, o fuzzing orientado por

inteligência artificial, a classificação de riscos com base em severidade, exploração conhecida e contexto operacional, e os mecanismos de documentação e monitoramento do ciclo de vida dos modelos não devem ser tratados como caminhos concorrentes, mas como camadas de uma mesma estratégia. Esse entendimento reforça a necessidade de arquiteturas híbridas, nas quais regras determinísticas, assinaturas, heurísticas, análise semântica de código, inferência estatística e validação humana se articulam em vez de disputar centralidade. A conclusão é clara: quanto mais crítico o ambiente corporativo, menos recomendável se torna a dependência exclusiva de um único método de detecção.

Ao mesmo tempo, este estudo evidencia que o entusiasmo com a inteligência artificial aplicada à cibersegurança deve ser acompanhado de rigor técnico, cautela metodológica e forte senso crítico. Embora a produção científica mostre avanços relevantes, persistem limitações estruturais que impedem generalizações excessivamente otimistas. Entre essas limitações, destacam-se a baixa qualidade de muitas bases de dados utilizadas em experimentação, o desbalanceamento entre classes, a fragilidade dos rótulos, a variação de desempenho entre domínios, a dificuldade de replicação dos resultados e a ainda insuficiente explicabilidade de muitos modelos. Tais restrições têm implicações práticas sérias. Em contextos organizacionais, um sistema que funciona bem em ambiente de teste, mas produz excesso de falsos positivos, baixa interpretabilidade ou degradação rápida de desempenho em produção, pode gerar mais ruído do que proteção. Portanto, outra conclusão relevante do estudo é que o sucesso da inteligência artificial em cibersegurança depende menos da promessa abstrata de alta acurácia e mais da sua utilidade concreta dentro de processos reais de triagem, correção, resposta e governança.

Essa constatação conduz a um ponto particularmente importante: a cibersegurança baseada em sistemas inteligentes não pode ser dissociada da qualidade da engenharia de software e da maturidade organizacional. Nenhum modelo, por mais sofisticado que seja, compensa integralmente uma cadeia de desenvolvimento insegura, ausência de inventário confiável de ativos, práticas fracas de gestão de configuração, inexistência de processos de atualização, negligência com dependências de terceiros ou baixa integração entre equipes de desenvolvimento, operações e segurança. A inteligência artificial pode melhorar a visibilidade e a priorização, mas não substitui controles fundamentais. Assim, a conclusão mais estratégica deste artigo é que os sistemas inteligentes devem ser concebidos como parte de um ecossistema maior de segurança, sustentado por desenvolvimento seguro, gestão contínua de vulnerabilidades, governança de risco, observabilidade operacional e resposta estruturada a incidentes. Em outras palavras, a inteligência só produz valor consistente quando se ancora em fundações institucionais maduras.

Outro resultado importante diz respeito à relevância dos referenciais técnicos e normativos na consolidação dessa área. Estruturas como CWE, CVSS, ATT&CK, catálogos de vulnerabilidades conhecidamente exploradas, software bill of materials, frameworks de desenvolvimento seguro e de gestão de risco cibernético fornecem vocabulário comum, critérios de classificação e mecanismos de interoperabilidade conceitual indispensáveis para que sistemas inteligentes não operem de forma isolada ou arbitrária. A inteligência computacional precisa ser alimentada por taxonomias confiáveis, parâmetros reconhecidos de severidade e contexto adversarial verificável. Sem essa base, corre-se o risco de transformar a detecção em mera geração automática de alertas sem

correspondência sólida com o risco real. Portanto, conclui-se que a inteligência em cibersegurança é, simultaneamente, algorítmica e normativa: depende tanto da capacidade de aprender padrões quanto da capacidade de alinhar-se a estruturas consolidadas de compreensão do risco.

Além disso, a pesquisa deixa evidente que a adoção de sistemas inteligentes traz consigo novas camadas de risco. Um dos pontos mais críticos discutidos foi o fato de que os próprios modelos podem ser alvo de manipulação, evasão, envenenamento de dados, roubo de parâmetros ou distorção do ambiente de inferência. Isso altera significativamente a forma como tais sistemas devem ser projetados e geridos. Já não basta avaliar o desempenho do modelo sobre um conjunto de dados; é preciso monitorar sua robustez diante de mudanças do ambiente, sua sensibilidade a entradas adversariais, a integridade das fontes de dados utilizadas em treinamento e a confiabilidade da cadeia tecnológica que o sustenta. Em consequência, a conclusão inevitável é que sistemas inteligentes aplicados à cibersegurança exigem uma segunda camada de segurança: devem ser protegidos, auditados, documentados e continuamente reavaliados. A defesa inteligente, portanto, precisa ser também uma defesa da própria inteligência.

No plano operacional, o estudo permite sustentar que a maior contribuição dos sistemas inteligentes em redes corporativas está na redução do tempo de exposição útil das vulnerabilidades. Em cenários reais, a dificuldade não está apenas em saber que uma vulnerabilidade existe, mas em determinar rapidamente se ela afeta ativos críticos, se há exploração ativa associada, qual sua prioridade relativa e que medida deve ser tomada com menor impacto operacional. Organizações frequentemente enfrentam excesso de

alertas, limitação de recursos humanos e filas de correções tecnicamente extensas. Nesses contextos, a aplicação de modelos inteligentes à priorização orientada por risco pode representar ganho substancial de efetividade. Não se trata apenas de rapidez técnica, mas de melhor alocação de atenção, energia organizacional e investimento defensivo. Assim, conclui-se que a inteligência artificial tende a ser mais valiosa quando aplicada à redução de ambiguidade decisória do que à simples multiplicação de diagnósticos.

Também se destacou, ao longo deste trabalho, a importância da explicabilidade, da documentação e da auditabilidade. Em cibersegurança, especialmente em ambientes corporativos e regulados, alertas e classificações precisam ser compreendidos por pessoas diferentes: analistas de segurança, desenvolvedores, gestores de tecnologia, auditores, responsáveis por conformidade e lideranças executivas. Um sistema que aponta risco, mas não explica minimamente por que o fez, dificulta a confiança, a validação e a ação corretiva. Por isso, esta pesquisa reforça que a evolução futura da área não depende apenas de melhores métricas de precisão, mas também de melhores mecanismos de interpretação e comunicação dos resultados. Modelos mais transparentes, relatórios mais contextualizados e práticas de documentação do ciclo de vida analítico serão cada vez mais centrais. A confiança operacional em sistemas inteligentes não nasce apenas do desempenho estatístico, mas da capacidade de o sistema se tornar inteligível e verificável dentro das rotinas institucionais.

A partir de tudo isso, pode-se afirmar que o futuro da detecção de vulnerabilidades em tempo real passa por uma convergência entre engenharia de software segura, inteligência artificial confiável,

automação defensiva e governança organizacional. Não haverá solução duradoura baseada apenas em tecnologia de detecção, assim como não haverá segurança suficiente baseada apenas em políticas formais sem visibilidade técnica contínua. O desafio contemporâneo consiste em articular esses elementos em uma arquitetura de defesa integrada, capaz de acompanhar a velocidade das mudanças tecnológicas e a criatividade das ameaças. Isso exige investimento em dados de qualidade, padronização de processos, observabilidade, capacitação de equipes, integração entre desenvolvimento e segurança, monitoramento contínuo de modelos e alinhamento com estruturas reconhecidas de risco e vulnerabilidade.

Em termos acadêmicos e científicos, esta pesquisa contribui para deslocar o debate de uma visão restrita e instrumental de inteligência artificial para uma visão sistêmica de sistemas inteligentes em cibersegurança. Ao analisar conjuntamente vulnerabilidade, exploração, monitoramento, risco, modelos analíticos e governança, o estudo oferece uma leitura mais compatível com a realidade organizacional. Em vez de tratar a inteligência artificial como fim, trata-a como componente de uma infraestrutura de decisão defensiva. Essa contribuição é relevante porque evita simplificações frequentes na literatura aplicada, nas quais resultados laboratoriais são confundidos com maturidade operacional. A pesquisa demonstra que a efetividade real depende de integração, contexto, curadoria de dados, documentação e capacidade humana de interpretação e resposta.

Por fim, conclui-se que os sistemas inteligentes para detecção de vulnerabilidades em tempo real possuem grande potencial transformador, mas esse potencial só será plenamente realizado se

acompanhado de responsabilidade técnica, clareza conceitual e compromisso institucional com a segurança como processo contínuo. O caminho mais promissor não está em prometer automação total ou substituição completa da expertise humana, mas em construir modelos híbridos, auditáveis e adaptativos, que ampliem a capacidade de observação e resposta sem comprometer transparência, governança e controle. Em um ambiente digital cada vez mais complexo, a cibersegurança eficaz dependerá menos de ferramentas isoladas e mais da capacidade de articular inteligência, processo e responsabilidade. Assim, a detecção em tempo real deve ser entendida como uma competência organizacional estratégica, sustentada por dados, engenharia e governança, capaz de transformar a segurança de um conjunto de reações tardias em um sistema contínuo de antecipação, priorização e resiliência.

REFERÊNCIAS BIBLIOGRÁFICAS

BILGIN, Zeki; ERSOY, Mehmet Akif; SOYKAN, Elif Üstündağ; TOMUR, Emrah; ÇOMAK, Pınar; KARAÇAY, Leyli. **Vulnerability Prediction From Source Code Using Machine Learning.** *IEEE Access*, 2020.

CHAFJIRI, Sadegh Bamohabbat; LEGG, Phil; HONG, Jun; TSOMPANAS, Michail-Antisthenis. **Vulnerability Detection Through Machine Learning-Based Fuzzing: A Systematic Review.** *Computers & Security*, 2024.

CISA. **Known Exploited Vulnerabilities Catalog.** 2026.

CISA. **Cybersecurity Performance Goals.** 2026.

FIRST. **Common Vulnerability Scoring System Version 4.0.** 2023-2026.

HARZEVILI, Nima Shiri; BELLE, Alvine Boaye; WANG, Junjie; WANG, Song; JIANG, Zhen Ming; NAGAPPAN, Nachiappan. **A Systematic Literature Review on Automated Software Vulnerability Detection Using Machine Learning.** *ACM Computing Surveys*, 2024.

MARJANOV, Tina; PASHCHENKO, Ivan; MASSACCI, Fabio. **Machine Learning for Source Code Vulnerability Detection: What Works and What Isn't There Yet.** *IEEE Security & Privacy*, 2022.

MITCHELL, Margaret; WU, Simone; ZALDIVAR, Andrew; BARNES, Parker; VASSERMAN, Lucy; HUTCHINSON, Ben; SPITZER, Elena; RAJI, Inioluwa Deborah; GEBRU, Timnit. **Model Cards for Model Reporting.** *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019.

MITRE. **Common Weakness Enumeration (CWE).** 2024-2026.

MITRE. **ATT&CK Framework.** 2026.

NIST. **The Cybersecurity Framework (CSF) 2.0.** 2024.

NIST. **Secure Software Development Framework (SSDF) Version 1.1.** 2022.

NIST. **Intrusion Detection and Prevention Systems.** 2007.

NIST. **Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations.** 2025.

NIST. **Software Bill of Materials (SBOM).** 2022.

NIST. **Static Analysis Tool Exposition (SATE/SAMATE).** 2010-2026.

OWASP. **Machine Learning Security Top 10.** 2023.

SEMASABA, Allan Oscar A.; YI, Z.; KHOMH, F.; FOUH, E. **An Empirical Evaluation of Deep Learning-Based Source Code Representations for Vulnerability Detection.** *Journal of Software: Evolution and Process*, 2022/2023.

---

[1] Doutor em Gestão de Projetos de Tecnologia da Informação e Doutorando em Engenharia Elétrica pela Universidade de Brasília (UnB). E-mail: romulodba@gmail.com

[2] Doutor em Ciências da Informação e em Altos Estudos Militares pelo Centro Universitário e Instituto de Educação Superior de Brasília (IESB). E-mail: paulo.borges@iesb.edu.br

[3] Especialista em Gestão e Inteligência em Segurança Pública, Tecnólogo em Gestão de Segurança Privada pelo Centro Universitário Internacional (UNINTER). E-mail: douglasestudosdas@gmail.com

[4] Mestre pela Universidade Federal de Viçosa (UFV). E-mail: marcus.d.souza@ufv.br

[5] Especialista em Engenharia Ambiental e Saneamento Básico, Graduado em Engenharia Elétrica pela Universidade Anhanguera Pitágoras Unopar. E-mail: cassionatanrl@hotmail.com