ESTUDO COMPARATIVO DE MODELOS SUPERVISIONADOS E NÃO SUPERVISIONADOS NA DETECÇÃO DE ANOMALIAS EM AMBIENTES DE INTERNET DAS COISAS (IOT): UMA REVISÃO DA LITERATURA

A COMPARATIVE STUDY OF SUPERVISED AND UNSUPERVISED MODELS FOR ANOMALY DETECTION IN INTERNET OF THINGS (IoT) ENVIRONMENTS: A LITERATURE REVIEW

Ana Iasmim Alves Maciel
Maria Clara Alves Soares
Pedro Medeiros Pitombeira
Tobias da Costa Gaspar
Karine Bessa Porto Pinheiro Vasques

RESUMO

O objetivo deste estudo é comparar aplicações de modelos supervisionados e não supervisionados na detecção de anomalias em ambientes IoT, identificando as técnicas mais eficazes. A justificativa está na necessidade de implementar defesas em sistemas IoT que acompanhem o nível das tecnologias contemporâneas, de modo que ataques que explorem essas tecnologias não se mostrem eficientes, dado o atual quadro de elevada vulnerabilidade enfrentado por tais sistemas. O método adotado foi uma revisão sistemática da literatura dentre o período mencionado e utilizando a base de dados Web of Science. Os resultados demonstraram que os modelos supervisionados, especialmente Random Forest e SVM, apresentam elevada precisão na identificação de ataques conhecidos, enquanto os modelos não supervisionados, como Isolation Forest e Autoencoders, destacam-se na detecção de ataques inéditos e anomalias sem a necessidade de dados rotulados. Além disso, observou-se uma tendência crescente para o uso de arquiteturas híbridas e soluções baseadas em TinyML e computação de borda. Conclui-se que não existe um modelo universalmente superior, sendo a escolha da técnica dependente das características do ambiente IoT, da disponibilidade de dados e das restrições computacionais da aplicação, embora abordagens híbridas se mostrem promissoras para o fortalecimento da segurança cibernética nesses sistemas.

Palavras Chaves: Segurança Cibernética. Inteligência artificial. Proteção de dados. Algoritmos de IA para sistemas IoT. Detecção de Anomalias. Aprendizado de Máquina. Aprendizado Profundo.

ABSTRACT

The objective of this study is to compare the application of supervised and unsupervised models in anomaly detection in IoT environments, identifying the most effective techniques. The rationale lies in the need to implement defenses in IoT systems that keep pace with contemporary technologies, so that attacks exploiting these technologies prove ineffective, given the current high level of vulnerability faced by such systems. The method adopted was a systematic literature review covering the specified period, using the Web of Science database. The results demonstrated that supervised models, particularly

Random Forest and SVM, exhibit high accuracy in identifying known attacks, while unsupervised models, such as Isolation Forest and Autoencoders, excel at detecting previously unseen attacks and anomalies without the need for labeled data. In addition, a growing trend toward the use of hybrid architectures and solutions based on TinyML and edge computing was observed. It is concluded that there is no universally superior model, and the choice of technique depends on the characteristics of the IoT environment, data availability, and the application's computational constraints, although hybrid approaches show promise for strengthening cybersecurity in these systems.

Keywords: Cybersecurity. Artificial intelligence. Data protection. AI algorithms for IoT systems. Anomaly detection. Machine learning. Deep learning.

1 INTRODUÇÃO

Nos últimos anos, houve um aumento exponencial na quantidade de dispositivos que utilizam tecnologia *Internet of Things* (IoT) , Esses ecossistemas, em geral, integram a *Edge Computing* (EC) para facilitar e automatizar várias tarefas diárias (Ahmim *et al.*, 2023) tanto no setor residencial quanto no industrial e comercial. O fenômeno IoT cresceu tão rapidamente que a projeção para 2025 era cerca de 75 bilhões de dispositivos conectados mundialmente (Vailshery, 2023 *apud* Zeng, 2024).

No entanto, ao mesmo tempo que a conexão com a internet promove uma série de vantagens aos aparelhos de EC, ela também sujeita os sistemas a ataques cibernéticos (Ahmim *et al.*, 2023). Além disso, a diversidade de dispositivos e a variedade de configurações existentes no sistema como um todo possibilitam ataques que afetam toda a rede através de um único ponto (Rehan; Demertzi, 2023 *apud* Zeng, 2024). Portanto, pode-se inferir que, quanto mais variadas forem as arquiteturas, maiores serão os problemas de segurança (Ahmim *et al.*, 2023).

Segundo Ali, Wang e Leung *et al.* (2025) *apud* Al-Quayed (2026), as vertentes da Inteligência Artificial(IA): Machine Learning(ML) e Deep Learning(DL), têm se mostrado pilares fundamentais da cibersegurança em dispositivos IoT. Em uma arquitetura IoT convencional, os dados captados pelos sensores são enviados para uma camada de nuvem; só então eles são analisados e uma inferência é realizada. No entanto, esse modelo centralizado resulta em alta latência e baixa qualidade de serviço (Samann *et al.*, 2023 *apud* Trilles; Hammad; Iskandaryan, 2024). Dessa maneira, torna-se necessário um modelo que contorne os desafios estabelecidos pela nuvem (Al-Quayed, 2026). Uma solução viável consiste em aplicar as funcionalidades antes executadas na nuvem diretamente em microcontroladores (MCUs), permitindo que o próprio dispositivo seja capaz de processar dados e realizar inferências de forma a detectar anomalias localmente. Essa abordagem arquitetural é denominada TinyML (Trilles; Hammad; Iskandaryan, 2024).

Porém, um desafio crítico desse modelo está justamente em seu núcleo: as abordagens tradicionais de Aprendizado de Máquina (ML) são incapazes de lidar perfeitamente com a volatilidade do tráfego real de dados. Isso ocorre devido à necessidade de dados históricos e concretos para estimar com precisão as anomalias, o que frequentemente resulta em falsos positivos ocasionados por desvios sazonais ou flutuações normais do próprio fluxo de dados. Embora pesquisas apontem modelos como *Random Forest* e *Support Vector Machine* (SVM) como os mais precisos, com acurácia superiores a 99% (Rafique *et al.*, 2024) a escassez de dados rotulados torna o desenvolvimento de um modelo de detecção de anomalias confiável baseado em aprendizado supervisionado um grande desafio (Trilles; Hammad; Iskandaryan, 2024). Logo, modelos não supervisionados são mais viáveis, pois dispensam a necessidade de rotulação prévia e são capazes de detectar ataques de dia zero — ou seja, ameaças inéditas

que não foram registradas no *dataset* de treinamento (Huda *et al.*, 2017; Mvula *et al.*, 2024 *apud* Al-Quayed, 2026).

Diante do que fora exposto, este trabalho tem por objetivo elencar trabalhos sobre o uso de modelos supervisionados e não supervisionados na detecção de anomalias em ambientes IoT, além de identificar as técnicas/algoritmos que apresentam maior desempenho por meio de uma análise comparativa dos trabalhos utilizados como base.

O presente artigo é uma revisão bibliográfica que está dividida da seguinte forma: o tópico 2 apresenta o referencial teórico que contextualiza os conceitos de detecção de anomalias na temática de aplicações dos modelos supervisionados e não supervisionados em ambientes IoT, o tópico 3 expõe a metodologia abordada no estudo, o tópico 4 traz a apresentação e análise dos dados obtidos e por fim, no tópico 5 tem discussões sobre os resultados obtidos, além das considerações finais sobre o tema.

2 REFERENCIAL TEÓRICO

2.1 Conceito de detecção de anomalias

A detecção de anomalias no ecossistema da Internet das Coisas (IoT) define-se como o processo de identificação de comportamentos que divergem significativamente dos padrões normais estabelecidos, sendo uma linha de defesa crítica em cenários onde soluções baseadas em assinaturas falham (Trilles; Hammad; Iskandaryan, 2024). De acordo com a literatura fundamental, anomalias são observações que levantam suspeitas de terem sido geradas por um mecanismo diferente ou que parecem inconsistentes com o restante do conjunto de dados (Hawkins, 1980 *apud* Trilles; Hammad; Iskandaryan, 2024). Tecnicamente, essas irregularidades são categorizadas em três variantes fundamentais para a análise de séries temporais em IoT: as pontuais, referentes a dados isolados fora do padrão; as contextuais, onde o dado é anômalo apenas dentro de um cenário ou período específico; e as coletivas, que representam sequências de dados que, agrupadas, indicam um comportamento irregular. A representação visual dessas distinções pode ser examinada na figura abaixo, que ilustra graficamente como esses desvios ocorrem ao longo do tempo.

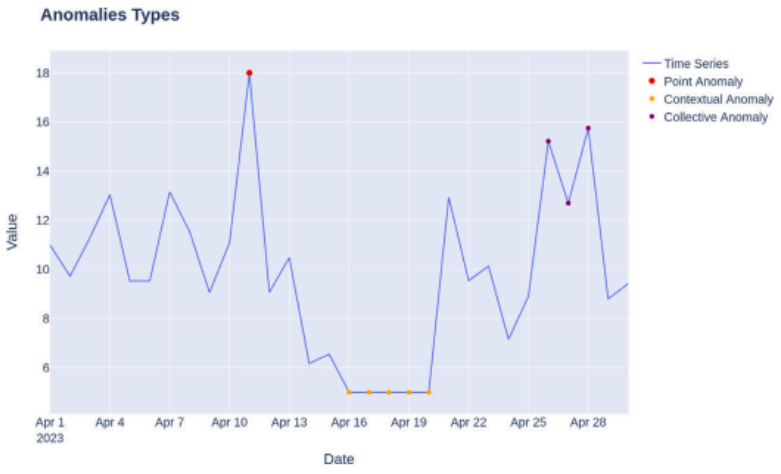

Fig. 3. Representation of the three existing types of anomalies.

Extraído de (Trilles; Hammad; Iskandaryan, 2024)

A necessidade urgente dessas abordagens decorre da dimensão gigantesca da "Internet de Tudo", que aumenta de forma expressiva a superfície de ataque em infraestruturas urbanas e em sistemas de saúde. (Zeng et al., 2024). Diferente dos métodos tradicionais, a Inteligência Artificial (IA) oferece uma abordagem proativa, sendo capaz de mimetizar o comportamento normal de um sistema e detectar ameaças emergentes, através de análise comportamental e aprendizado adaptativo (Al-Quayed, 2026; Zeng *et al.,* 2024)

2.2 Modelos Supervisionados

Os modelos de aprendizado supervisionado operam através da predição de rótulos baseados em classes previamente definidas durante o treinamento, sendo amplamente reconhecidos por sua precisão na classificação de ataques conhecidos (Trilles; Hammad; Iskandaryan, 2024). Nos artigos, o algoritmo Random Forest (RF) é consistentemente citado como um dos mais eficazes para sistemas de detecção de intrusão, alcançando taxas de acurácia superiores a 99% em diversos estudos (Rafique *et al*., 2024). Além do RF, as Redes Neurais Convolucionais (CNN) destacam-se pela capacidade de extrair características espaciais complexas de tráfego de rede e dados de sensores, sendo o modelo supervisionado mais frequente em aplicações de TinyML devido à sua eficiência em dispositivos de recursos limitados (Trilles; Hammad; Iskandaryan, 2024).

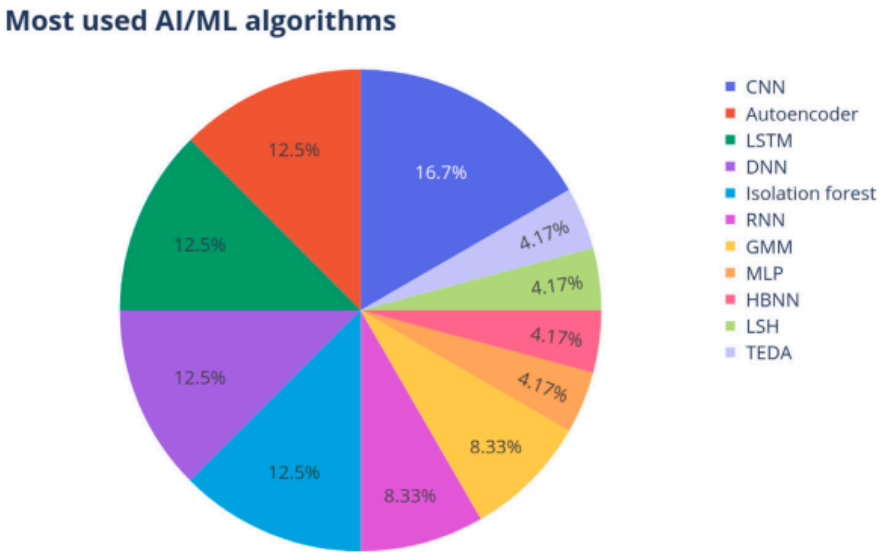

Fig. 9. Distribution of articles by ML/DL algorithms.

Extraído de (Trilles; Hammad; Iskandaryan, 2024)

Entretanto, a dependência de dados rotulados representa um desafio técnico significativo, pois, em ambientes reais de IoT, a rotulagem de vastos volumes de dados de rede é custosa e muitas vezes inviável (Trilles; Hammad; Iskandaryan, 2024). Para mitigar esses riscos e garantir o desempenho, os pesquisadores utilizam métricas rigorosas como Acurácia, F1-Score e matrizes de confusão para validar a confiabilidade dos classificadores em cenários de alta criticidade (Rafique *et al.*, 2024).

2.3 Modelos Não Supervisionados

As abordagens de aprendizado não supervisionado são essenciais para a cibersegurança em IoT por sua habilidade de aprender a estrutura intrínseca dos dados sem a necessidade de rótulos prévios, tornando-as a principal defesa contra ameaças inéditas (Al-Quayed, 2026). Entre as técnicas de destaque estão os Autoencoders, que identificam anomalias através do erro de reconstrução de dados, e o Isolation Forest, que foca em isolar instâncias anômalas em vez de perfilar o comportamento normal, sendo altamente eficiente para grandes volumes de dados (Trilles; Hammad; Iskandaryan, 2024). A tendência atual do estado da arte reside na hibridização, que combina diferentes tipos de redes neurais profundas para explorar propriedades distintas e alcançar alta performance (Ahmim *et al*., 2023). Um exemplo técnico dessa integração paralela, que une CNNs e LSTMs com Autoencoders para extrair características ocultas e classificar ataques complexos, é detalhado na figura abaixo.

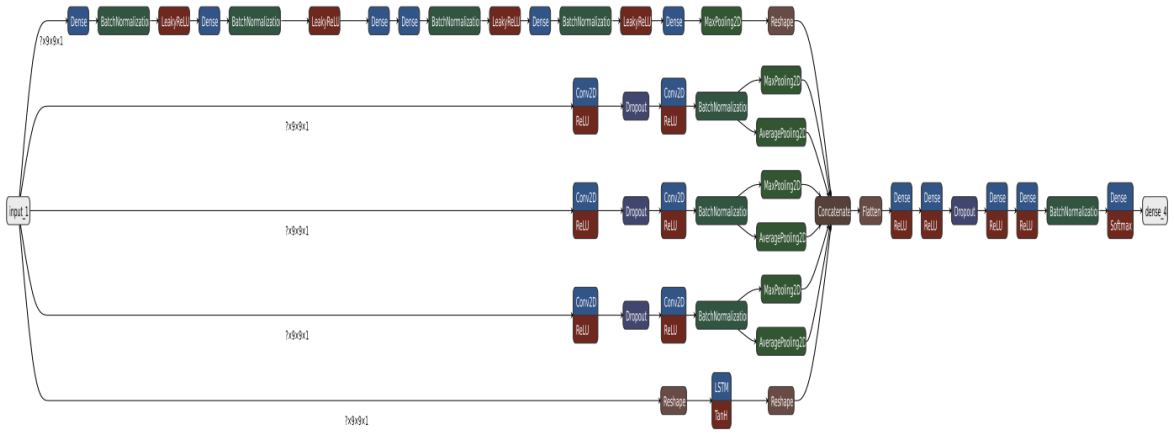

FIGURE 3. Architecture of the sub model.

Extraído de (Ahmim *et al.*, 2023)

Além da eficácia na detecção, a descentralização é vital para aplicações sensíveis ao tempo. Estudos indicam que a distribuição inteligente de modelos pode reduzir a latência em modelos centralizados na nuvem de 152,3ms para apenas 18,7ms na borda da rede (Al-Quayed, 2026).

3 METODOLOGIA

Considerando o contexto estabelecido nesta pesquisa e tendo em vista o propósito central de analisar comparativamente, classificar e categorizar a aplicação de modelos supervisionados e não supervisionados no mapeamento e na detecção de comportamentos anômalos em ecossistemas de Internet das Coisas (IoT), estabelecem-se as diretrizes procedimentais a seguir.

A condução desta revisão da literatura baseou-se nos fundamentos metodológicos propostos por Souza e Santos (2020), o qual foi devidamente adaptado para fazer frente às especificidades técnicas do conjunto de redes e inteligência computacional. Este modelo metodológico estrutura-se sobre três macro etapas sequenciais e interdependentes: o Planejamento da Revisão, a Condução da Pesquisa e, por fim, a Divulgação dos Resultados.

Para assegurar um rastreamento operacional transparente e cuidadoso do estudo, essas macro fases foram divididas em seis micro etapas de execução prática, as quais encontram-se mapeadas e descritas em termos de atividades na Tabela 1.

Tabela 1 - Etapas da Pesquisa

Macro etapas da Revisão	**Micro etapas Operacionais**	**Atividades e Escopo Desenvolvidos**
Planejamento	Etapa 1	Definição do tema, do problema de pesquisa e dos objetivos gerais e específicos.
Condução	Etapa 2	Seleção e validação da base de dados internacional para a coleta de informações.
	Etapa 3	Definição do período da busca para captar tendências tecnológicas atuais.
	Etapa 4	Formulação das buscas com palavras-chave e operadores booleanos.
	Etapa 5	Definição e aplicação dos critérios de inclusão e exclusão dos artigos.
Divulgação dos Resultados	Etapa 6	Sistematização, leitura analítica e consolidação dos dados extraídos do conjunto de artigos selecionados.

Fonte: Elaborado pelos autores (2026) a partir de Souza e Santos (2020).

A primeira etapa do trabalho consistiu na fase de Planejamento, fase em que ocorreu a especificação e delimitação temática da pesquisa. A escolha desse objeto de estudo justificou-se pelo considerável avanço recente no volume de dispositivos baseados em Internet das Coisas (IoT) operando de forma integrada à computação de borda (*Edge Computing*). Embora a descentralização melhore de forma significativa a latência e a qualidade do serviço ao processar dados localmente, ela fragmenta o perímetro de segurança tradicional.

A diversidade de fabricantes, a diversidade das configurações de hardware e a volatilidade do tráfego real de rede criam ambientes propensos a invasões furtivas, onde o comprometimento de um único ponto pode paralisar toda a malha de dispositivos conectados. Esse cenário de hiperconectividade gerou um panorama de vulnerabilidade acentuada contra incidentes virtuais, exigindo o desenvolvimento e a implementação de mecanismos de defesa cibernética

robustos, como o paradigma *TinyML*, capazes de executar inferências diretamente em microcontroladores limitados para identificar ameaças em tempo real.

Na fase seguinte, voltada à Condução da pesquisa, determinou-se o repertório científico digital para a execução das buscas automatizadas. Optou-se pela utilização exclusiva da base de dados internacional *Web of Science*. A escolha dessa plataforma se deve ao seu amplo reconhecimento na comunidade acadêmica internacional, a qual é conhecida por indexar apenas periódicos de alta qualidade, com avaliação rigorosa por pares e cobertura de diversas áreas do conhecimento. Essa seleção prévia da base aumenta a confiabilidade da revisão, garantindo que os estudos e dados analisados apresentem consistência metodológica e estatística.

O intervalo temporal da pesquisa foi delimitado com o intuito de consolidar as investigações mais recentes do campo e mapear as disrupções tecnológicas diretamente associadas ao surgimento da Indústria 4.0 e das redes inteligentes. Este intervalo temporal reflete o momento de transição em que as abordagens convencionais de segurança, baseadas puramente em assinaturas estáticas de malwares, começaram a falhar diante da instabilidade e complexidade dos ataques cibernéticos modernos.

Foi nesse cenário que as técnicas de inteligência artificial aplicada ganharam crescimento comercial e acadêmica, desafiando as restrições físicas de hardware e exigindo que modelos matemáticos complexos fossem otimizados para rodar em camadas de borda (*Edge*) e névoa (*Fog*). Desse modo, o período selecionado garante que o diagnóstico gerado por esta revisão reflita as práticas de engenharia de software, as métricas de desempenho e os desafios computacionais do cenário atual.

A definição da estratégia de busca e a seleção coesa das palavras-chave foram passos importantes para assegurar que a busca na base encontrasse apenas os estudos mais relevantes para o tema, evitando tanto a perda de informações relevantes quanto a inclusão de conteúdos pouco pertinentes. A estratégia de busca utilizou termos internacionais organizados em três eixos: segurança cibernética ("Segurança Cibernética", "Proteção de Dados" e "Detecção de Anomalias"), inteligência artificial ("Inteligência Artificial", "Aprendizado de Máquina" e "Deep Learning") e aplicações em IoT ("Algoritmos de IA para Sistemas IoT"). A combinação desses termos permitiu identificar estudos sobre o uso de algoritmos de IA na proteção e gestão de redes conectadas.

A quinta etapa metodológica envolveu a definição dos filtros de elegibilidade, com critérios de inclusão e exclusão aplicados aos estudos encontrados na busca inicial. Para qualificar o portfólio final e garantir comparações justas de desempenho, os estudos recuperados precisavam atender aos seguintes critérios, sob pena de exclusão:

a. Estar contido dentro do intervalo temporal estabelecido para a revisão da literatura;
b. Apresentar relação com abordagens de detecção de anomalias em redes ou dispositivos de IoT;
c. Possuir link ativo para download do texto completo, permitindo acesso e extração dos dados.;
d. Não consistir em registro duplicado ou versão preliminar na base de dados;
e. Abordar a modelagem, o treinamento ou a validação de pelo menos uma técnica de Machine Learning ou Deep Learning;
f. Fornecer dados quantitativos sobre o desempenho dos detectores, incluindo métricas para modelos supervisionados e indicadores de qualidade para modelos não supervisionados.

A pesquisa realizada na plataforma *Web of Science* resultou em um conjunto inicial de artigos científicos pré-selecionados. Posteriormente, todos os estudos passaram por uma triagem qualitativa individual, com análise completa dos textos e comparação com as necessidades de sistemas IoT. Foram descartados estudos que abordavam segurança de forma genérica ou que propunham modelos voltados apenas para nuvem, sem considerar limitações de energia, memória e latência em dispositivos embarcados.

Após a aplicação dos critérios de filtragem, foi definida a amostra final de artigos. Esse conjunto serviu como base teórica e prática da revisão. A partir dele, foram extraídas informações sobre métodos, parâmetros, limitações e cenários de ataque dos algoritmos, organizados em duas abordagens: supervisionada e não supervisionada.

Nos modelos supervisionados, a análise priorizou métricas da matriz de confusão (acurácia, precisão, recall e F1-score), evitando depender apenas da acurácia devido ao forte desbalanceamento dos dados. O objetivo foi verificar a capacidade dos modelos de reduzir falsos negativos sem comprometer a eficiência em ambientes com restrições computacionais. Já nos modelos não supervisionados, como métodos de agrupamento e detecção de anomalias, a avaliação exigiu abordagens alternativas, pois esses algoritmos trabalham com dados sem rótulos e não permitem validação direta por tabelas de desempenho.

Por fim, na última etapa os dados dos artigos selecionados foram organizados e analisados criticamente, reunindo resultados sobre técnicas de Machine Learning e Deep Learning, assim como os diagramas, as discussões analíticas a respeito de seus desempenhos comparativos e as considerações finais sobre suas aplicações em segurança cibernética em ambientes de IoT.

3.1 Diretrizes sobre o Uso de Inteligência Artificial

No desenvolvimento deste artigo, os autores utilizaram as ferramentas de Inteligência Artificial Generativa a seguir como recurso auxiliar para revisão linguística, organização estrutural de ideias e aprimoramento da clareza expositiva do texto: Gemini e NotebookLM, desenvolvido pela Google (2026). Todo o conteúdo científico — incluindo a seleção das referências, a análise comparativa dos algoritmos, a interpretação dos resultados e as conclusões apresentadas — foi inteiramente concebido, verificado e validado pelos autores, que assumem integral responsabilidade pela precisão, integridade e originalidade das informações contidas neste trabalho. O uso de IA não substituiu o julgamento crítico dos pesquisadores em nenhuma etapa do processo investigativo.

4 APRESENTAÇÃO DOS DADOS

4.1 Conceito de Detecção de Anomalias

No contexto da segurança de sistemas IoT, a detecção de anomalias consiste no mecanismo responsável por identificar instâncias em que o comportamento de um sistema se desvia do padrão estabelecido como normal, viabilizando a identificação e a resolução de irregularidades antes que causem danos ao ecossistema (Rafique *et al*., 2024). Quando uma instância de dados apresenta um comportamento diferente do esperado pelo sistema, ela passa a ser denominada anomalia — também referida na literatura de mineração de dados como anormalidade, novidade, desvio ou discrepância (Maia; Lemos, 2021). A combinação de Inteligência Artificial com ecossistemas IoT potencializa a eficácia desse processo, melhorando a confiabilidade dos sistemas, acelerando os tempos de resposta e viabilizando a detecção em tempo real (Rafique *et al.*, 2024; Al-Quayed, 2026). Sistemas inteligentes de detecção de intrusões baseados em IA são capazes de identificar uma ampla gama de

ameaças, incluindo ataques de força bruta, inundação de requisições, injeção de comandos, negação de serviço distribuída (DDoS) e exploração de backdoors (Ahmim *et al.*, 2023; Rafique et al., 2024). A literatura classifica as anomalias em três categorias fundamentais: pontuais, contextuais e coletivas, cada uma com características estruturais e desafios de detecção distintos (Trilles; Hammad; Iskandaryan, 2024; Maia; Lemos, 2021).

Tipo de Anomalia	**Característica Principal**	**Exemplo em IoT**	**Algoritmos Indicados**
Pontual	Observação isolada desviante	Pico abrupto de tráfego em um sensor	Isolation Forest, k-Means
Contextual	Desvio dependente do contexto temporal/espacial	Leitura atípica fora do horário de operação	LOF, Autoencoders
Coletiva	Conjunto de instâncias normais que forma padrão anômalo	Ataque DDoS distribuído entre dispositivos	GMM, Agglomerative Clustering

Fonte: Elaborado pelos autores (2026), adaptado a partir de Trilles, Hammad e Iskandaryan (2024), Rafique et al. (2024), Maia e Lemos (2021).

A distinção entre os três tipos de anomalia é determinante para a seleção do algoritmo mais adequado ao perfil de ameaças de um dado ecossistema IoT. As anomalias pontuais, por serem o tipo mais elementar e o foco da maioria das pesquisas, são bem atendidas por algoritmos como Isolation Forest e k-Means Clustering. Já as anomalias contextuais e coletivas demandam modelos com maior capacidade de representação temporal e estrutural, como Autoencoders e Gaussian Mixture Models (Rafique et *al.,* 2024). Ambientes críticos, como plantas industriais da Indústria 4.0 e infraestruturas de saúde, frequentemente demandam a implementação combinada de detectores para os três tipos, dado que diferentes vetores de ataque exploram diferentes dimensões do comportamento do sistema (Al-Quayed, 2026; Zeng *et al.*, 2024).

4.2 Modelos Supervisionados

A engenharia por trás do Aprendizado de Máquina Supervisionado opera sob uma lógica baseada no treinamento de sistemas preditivos a partir de históricos de dados

previamente classificados por rótulos conhecidos. No ecossistema da Internet das Coisas (IoT), esse método consolidou-se como uma linha de defesa cibernética essencial, conforme apontam Rafique *et al*. (2024). O motivo é essencialmente prático: ao lidar com a segurança de infraestruturas conectadas, os classificadores supervisionados entregam alta precisão ao separar o tráfego legítimo de anomalias que já possuem assinaturas mapeadas na literatura (Zeng *et al.*, 2024).

O motor que move esse paradigma é o aprendizado indutivo estruturado com base em exemplos rotulados. O algoritmo supervisionado analisa o comportamento estatístico das variáveis de rede e monta uma fronteira de decisão direcionada. Quando um pacote de dados inédito atinge o gateway (porta de entrada) ou a camada de gerenciamento, o classificador valida a instância de acordo com os padrões assinados na fase de treino (Trilles; Hammad; Iskandaryan, 2024). Essa abordagem baseada em modelos supervisionados funciona como um escudo direcionado contra ameaças mapeadas.

A implementação de mecanismos supervisionados nos cenários de IoT impõe considerações profundas sobre a eficiência energética e a capacidade de memória dos dispositivos (Al Quayed, 2026). Essas ferramentas precisam ser otimizadas para atuar nas camadas de borda (Edge) e névoa (Fog), blindando ambientes dinâmicos que geram fluxos massivos de dados, como redes médicas e plantas industriais modernas ligadas ao ecossistema da Indústria 4.0, sem causar atrasos prejudiciais na transmissão das mensagens (Rafique et al., 2024; Souza; Santos, 2020).

4.2.1 Classificadores Supervisionados Aplicados à Triagem de Anomalias

A seleção de um algoritmo supervisionado depende do tipo de dado capturado pelos sensores, da natureza das ameaças e das limitações computacionais do hardware em campo. Como a eficácia de um modelo supervisionado nasce do alinhamento entre as propriedades do tráfego, as restrições da arquitetura física e a estrutura do classificador, a literatura de IoT mapeia diferentes abordagens para a mitigação de ataques (Ahmim *et al*., 2023; Rafique *et al*., 2024).

O Quadro 1 detalha os principais algoritmos de modelagem supervisionada identificados nos mapeamentos da literatura de IoT, correlacionando seus mecanismos, limites lógicos e vetores de ataque alvo:

Quadro 1 Dados de Descrição Avançada dos Algoritmos Supervisionados Aplicados à IoT

Algoritmo	**Mecanismo de Classificação Supervisionada em Redes IoT**	**Parâmetros de Entrada e Limitações Estruturais**	**Impacto Prático na Segurança e Ataques Alvo**
Decision Tree	Cria uma árvore lógica supervisionada de perguntas e respostas. O modelo analisa as variáveis do tráfego e ramifica as decisões até as folhas para definir se o fluxo representa tráfego normal ou ataque (Rafique *et al*., 2024).	Entradas: Endereços de IP, portas lógicas e tamanho do pacote. Limitação: Propensa ao sobreajuste se a profundidade da árvore crescer sem limites.	Ataques Alvo: Invasões por força bruta e varreduras de portas (Port Scanning). Impacto: Garante decisões rápidas em gateways locais.
Support Vector Machine (SVM)	Busca uma superfície de separação ideal no espaço de dados chamada hiperplano. Usa as amostras mais difíceis da borda como vetores de suporte para maximizar a distância entre as classes (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Vetores numéricos normalizados de consumo de banda e frequência de conexões. Limitação: O custo de processamento cresce de forma quadrática com o volume de dados.	Ataques Alvo: Anomalias em payloads e injeção de comandos falsos. Impacto: Alta precisão em fronteiras de dados bem definidas.

Random Forest	Constrói um comitê composto por múltiplas árvores de decisão supervisionadas. Cada estrutura analisa uma fração sorteada dos logs da rede e o veredito final é definido por voto majoritário (Ahmim *et al*., 2023).	Entradas: Matrizes densas de telemetria de sensores e logs de pacotes brutos. Limitação: Exige alta capacidade de memória RAM para armazenar as árvores simultaneamente.	Ataques Alvo: Ataques distribuídos de negação de serviço (DDoS) complexos. Impacto: Extremamente estável frente a logs ruidosos ou falhas de sensores.
MultiLayer Perceptron (MLP)	Rede neural clássica supervisionada organizada em camadas que ajusta os pesos das suas conexões internas por retropropagação, capturando padrões de intrusão complexos e não lineares (Rafique *et al*., 2024).	Entradas: Características extraídas e pré processadas do fluxo de tráfego líquido. Limitação: Fase de treinamento lenta e dependência de grandes volumes de instâncias.	Ataques Alvo: Tentativas de exfiltração de dados camufladas no tráfego. Impacto: Excelente capacidade de generalização para novos perfis correlacionados.

AdaBoost	Método de treinamento supervisionado em série que cria classificadores fracos sequencialmente, aumentando o peso sobre os erros cometidos na etapa anterior para se especializar em fluxos difíceis (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Logs de tráfego onde as classes de ataque estão sub-representadas. Limitação: Sensível a ruídos e dados discrepantes que descalibram os pesos sequenciais.	Ataques Alvo: Ataques furtivos e persistent de baixa frequência. Impacto: Maximiza o acerto em amostras que outros modelos ignoram.
k Nearest Neighbor (k NN)	Classifica novos pacotes por proximidade geométrica dentro do mapa de características rotuladas, avaliando os k registros salvos mais perto da nova amostra para tomar a decisão (Rafique *et al*., 2024).	Entradas: Estatísticas de fluxo temporal de pacotes através de métricas de distância. Limitação: Alto tempo de execução na fase de teste devido ao cálculo contínuo de distâncias.	Ataques Alvo: Desvios abruptos de comportamento em sensores industriais. Impacto: Eficaz em bases onde o comportamento normal é muito coeso.

Naive Bayes	Trabalha com estatística de probabilidades baseada no Teorema de Bayes, partindo do princípio de que os atributos da rede não dependem uns dos outros, o que gera um processamento rápido em chips limitados (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Frequência de ocorrência de comandos e strings em payloads de protocolos IoT. Limitação: A premissa de independência total entre as variáveis raramente ocorre na realidade.	Ataques Alvo: Triagem inicial de malwares baseada em palavras chave de pacotes. Impacto: Ideal para implementação direta em sensores com bateria limitada.
Extra Trees	Funciona de modo semelhante à Random Forest, mas define os cortes e divisões das suas árvores de maneira totalmente aleatória, economizando tempo de processamento em bases volumosas (Rafique *et al*., 2024).	Entradas: Grandes volumes de dados estatísticos de fluxos de redes heterogêneas. Limitação: Pode gerar caminhos lógicos menos interpretáveis do que uma árvore convencional.	Ataques Alvo: Inundação de requisições de rede (Flooding). Impacto: Processamento ultra veloz para ambientes de Big Data em IoT.

Gradient Boosting	Monta árvores de decisão de forma sequencial otimizando o sistema supervisionado por meio de descida de gradiente, focando estritamente na correção dos erros deixados pelas etapas anteriores (Ahmim *et al*., 2023).	Entradas: Métricas de variância de jitter, latência e taxa de erro de pacotes. Limitação: Exige um ajuste complexo de hiperparâmetros para evitar o sobreajuste rápido.	Ataques Alvo: Ataques de negação de serviço em camadas de aplicação. Impacto: Entrega um dos maiores índices de F1 Score da literatura técnica.
Neural Networks	Modelos robustos de aprendizado profundo supervisionado que ajustam parâmetros de forma autônoma, rastreando se um fluxo faz parte de campanhas complexas de invasão (Ahmim *et al*., 2023; Rafique *et al*., 2024).	Entradas: Sequências temporais brutas de pacotes IP e fluxos contínuos de tráfego. Limitação: Opacidade analítica que dificulta a auditoria direta dos alertas gerados.	Ataques Alvo: Sequestro de dispositivos (Botnets) e tráfego malicioso encriptado. Impacto: Elimina a necessidade de extração manual de características.

Logistic Regression	Algoritmo estatístico supervisionado clássico que calcula a chance de um evento ser verdadeiro ou falso dentro de uma curva logística, sendo uma ferramenta leve para sensores pequenos (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Variáveis binárias ou contínuas de pacotes de comunicação simplificados. Limitação: Incapaz de resolver correlações complexas e não lineares de segurança.	Ataques Alvo: Spoofing de identidade de hardware básico. Impacto: Consumo computacional quase nulo na camada de borda.
TabNet	Arquitetura de rede neural pensada para ler tabelas de logs estruturados que aplica mechanisms de atenção para focar nas colunas que importam para a segurança cibernética (Rafique *et al*., 2024).	Entradas: Tabelas estruturadas de NetFlow contendo múltiplas colunas e métricas de rede. Limitação: Alto custo de processamento durante a fase de treino inicial.	Ataques Alvo: Manipulação de registros históricos e anomalias tabulares. Impacto: Oferece interpretabilidade nativa mantendo o poder de Deep Learning.

Linear Discriminant Analysis (LDA)	Projeta os dados buscando uma linha de separação que maximize a distância entre a média do tráfego limpo e a média dos ataques, consumindo pouca memória (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Distribuições de tráfego contínuas que seguem uma distribuição gaussiana. Limitação: Perde eficácia se as matrizes de covariância das classes forem muito diferentes.	Ataques Alvo: Intrusões com perfis de tráfego geometricamente lineares. Impacto: Baixíssima latência na tomada de decisão preditiva.
Quadratic Discriminant Analysis (QDA)	Evolução do LDA que aceita que as classes supervisionadas tenham comportamentos dispersos e formatos diferentes na rede, desenhando curvas de decisão quadráticas (Trilles; Hammad; Iskandaryan, 2024).	Entradas: Logs de sensores industriais com comportamento multimodal ou disperso. Limitação: Requer a estimação de um número maior de parâmetros estatísticos.	Ataques Alvo: Desvios complexos em barramentos de automação. Impacto: Captura variações estatísticas sutis sem exigir redes neurais.

LightGBM	Framework de Gradient Boosting focado no crescimento de árvores olhando para os nós das folhas mais profundas, projetado para devorar bancos de dados volumosos com baixo uso de memória RAM (Ahmim *et al*., 2023).	Entradas: Conjuntos massivos e distribuídos de logs históricos de ataques de rede. Limitação: Pode causar sobreajuste em bases de dados pequenas com poucos registros.	Ataques Alvo: Ataques volumétricos de negação de serviço em tempo real. Impacto: Reduz dramaticamente o tempo de resposta em ambientes de nuvem.

Fonte: Elaborado pelos autores (2026), adaptado a partir de Al Quayed (2026), Zeng et al. (2024), Ahmim et al. (2023), Trilles, Hammad e Iskandaryan (2024) e Rafique et al. (2024).

4.2.2 Métricas de Avaliação Computacional para Detetores Rotulados

Colocar um classificador supervisionado à prova exige fatiar a base de dados rotulada em subconjuntos distintos. O desenho padrão documentado na literatura envolve abastecer a maior fatia do banco para o treinamento, permitindo que o modelo calibre seus pesos, enquanto o pedaço restante será o grupo de teste, cuja função é simular o mundo real e testar o poder de generalização do algoritmo frente a dados inéditos (Ahmim *et al*., 2023; Rafique *et al*., 2024).

A régua usada para medir o sucesso dessa classificação é a Matriz de Confusão, que cruza os alertas dados pelo modelo com os rótulos originais, separando a performance em quatro cenários (Trilles; Hammad; Iskandaryan, 2024):

- Verdadeiros Positivos (TP): Invasões e ataques reais identificados de forma correta.
- Verdadeiros Negativos (TN): Fluxos de dados normais que o sistema identificou como legítimos.
- Falsos Positivos (FP): Os alarmes falsos; comunicações limpas julgadas como perigosas.
- Falsos Negativos (FN): A falha mais crítica, ocorrendo quando um ataque consegue se camuflar e o modelo o classifica como tráfego normal (Ahmim *et al*., 2023).

A partir desse mapeamento, extraem-se os cálculos de avaliação do modelo supervisionado:

Acurácia

Porcentagem geral de acertos do classificador supervisionado, dividindo tudo o que ele acertou pelo volume total de testes:

$$\text{Acurácia} = (TP + TN) / (TP + FP + TN + FN)$$

Precisão
Mede o nível de confiança dos alertas disparados. De todos os alarmes de anomalia emitidos, quantos eram ataques reais:

Precisão = TP / (TP + FP)

Revocação (Recall)
Avalia a capacidade de varredura do sistema. De todos os ataques que tentaram invadir o ecossistema IoT, qual a fração que o classificador conseguiu capturar:

Recall = TP / (TP + FN)

F1 Score
Média harmônica que pondera o equilíbrio entre a Precisão e o Recall, servindo como a métrica mais confiável quando há assimetria de impacto na segurança (Rafique *et al*., 2024):

F1 Score = 2 * (Precisão * Recall) / (Precisão + Recall)

Existe uma particularidade no tráfego de redes IoT que desafia os modelos supervisionados: o desbalanceamento massivo de classes. Em uma infraestrutura real, mais de 99% das mensagens correspondem a rotinas limpas (Ahmim *et al.*, 2023; Trilles; Hammad; Iskandaryan, 2024). Por conta dessa desproporção, a literatura faz um alerta enfático: olhar para a Acurácia de forma isolada é um erro. Se um modelo classificar todas as mensagens como normais por viés, ele baterá uma acurácia de 99%, mas sua utilidade será nula por deixar passar 100% das invasões. A análise aqui confronta esses números com o F1 Score e com a capacidade de manter os Falsos Negativos perto de zero em ambientes complexos de IoT.

4.3 Modelos Não Supervisionados

Diferente da abordagem baseada em rótulos preexistentes, o Aprendizado de Máquina Não Supervisionado opera sem a necessidade de uma verdade absoluta para categorizar os registros do tráfego. No contexto da Internet das Coisas, onde novas assinaturas de ataques surgem diariamente e os dispositivos geram um fluxo contínuo de dados brutos, esse paradigma destaca se pela habilidade de mapear o desconhecido e identificar ataques de dia zero (Zeng *et al*., 2024).

A arquitetura lógica desse modelo fundamenta se na análise estrutural e geométrica das variáveis. O sistema assume que o tráfego legítimo de um ecossistema de sensores tende a seguir um padrão comportamental previsível e repetitivo. Consequentemente, qualquer pacote que exiba um desvio estatístico relevante em relação ao perfil consolidado é isolado como uma anomalia potencial (Trilles; Hammad; Iskandaryan, 2024). Essa metodologia elimina a dependência de bancos de assinaturas estáticos que frequentemente falham diante de campanhas de invasão modernas.

Contudo, a implementação de mecanismos não supervisionados nos cenários de IoT impõe restrições de processamento e armazenamento (Al Quayed, 2026; Rafique *et al*., 2024). Os algoritmos precisam processar dados multidimensionais vindos de redes industriais da Indústria 4.0 em tempo real, operando muitas vezes de forma distribuída entre a névoa e a borda da rede. O desafio analítico reside em construir fronteiras de agrupamento robustas para mitigar os alarmes falsos sem sobrecarregar o hardware embarcado (Rafique *et al*., 2024; Souza; Santos, 2020).

4.3.1 Classificadores Não Supervisionados Aplicados à Detecção de Anomalias

A seleção do classificador não supervisionado é condicionada pelas propriedades estatísticas das variáveis e pela topologia da rede IoT. O sucesso de um modelo heurístico depende do ajuste geométrico entre as premissas do algoritmo e a distribuição espacial dos logs de tráfego capturados (Ahmim *et al*., 2023; Rafique *et al*., 2024).

O Quadro 2 apresenta as principais técnicas não supervisionadas documentadas nos mapeamentos de segurança para IoT, discriminando suas abordagens algorítmicas e comportamento matemático diante de dados desconhecidos:

Quadro 2 Dados de Descrição Avançada dos Algoritmos Não Supervisionados Aplicados à IoT

Algoritmo	**Mecanismo de Classificação Não Supervisionada em Redes IoT**	**Estratégia de Agrupamento e Limitações Operacionais**	**Comportamento Frente ao Desconhecido e Anomalias**
k Means Clustering	Agrupa os logs dividindo os dados em k clusters com base na distância euclidiana até os centroides. Identifica anomalias calculando registros que se posicionam muito distantes do centro dos grupos normais (Rafique *et al*., 2024).	Estratégia: Minimização da variância interna dos grupos com base em centroides fixos. Limitação: Exige que o usuário defina previamente o valor exato de k.	Ação: Identifica picos anormais isolados fora dos grupos de rotina estruturados. Foco: Mapeamento de desvios grosseiros de comportamento.

Isolation Forest	Isola anomalias de forma explícita por meio de árvores de decisão aleatórias. Como os ataques possuem características discrepantes, necessitam de menos partições lógicas para serem isolados nas folhas (Trilles; Hammad; Iskandaryan, 2024).	Estratégia: Particionamento recursivo do espaço amostral através de cortes aleatórios nas variáveis. Limitação: Dificuldade para identificar anomalias se estas estiverem muito concentradas.	Ação: Captura anomalias globais de forma veloz isolando pacotes suspeitos. Foco: Altamente eficiente contra ataques de dia zero em fluxos massivos.
Local Outlier Factor (LOF)	Calcula a densidade local de uma amostra de rede em comparação com a vizinhança. Fluxos com densidade substancialmente inferior à de seus vizinhos são classificados como desvios ou intrusões (Rafique *et al*., 2024).	Estratégia: Comparação da densidade local de um ponto com a densidade dos seus k vizinhos mais próximos. Limitação: O cálculo consome muita memória quando aplicado a fluxos contínuos.	Ação: Detecta anomalias locais ocultas dentro de subestruturas de dados densos. Foco: Captura ataques que imitam em parte o comportamento legítimo.

One Class SVM	Projetada para cenários onde apenas dados do tráfego legítimo estão disponíveis no treino. O algoritmo desenha um limite geométrico firme em torno dos dados normais, barrando o que ficar de fora (Trilles; Hammad; Iskandaryan, 2024).	Estratégia: Mapeamento dos dados normais para um espaço de alta dimensão usando funções de kernel. Limitação: Desempenho afetado se existirem dados de ataque camuflados no treino.	Ação: Cria um envelope de segurança rígido tolerando apenas o comportamento homologado. Foco: Blindagem estrita contra qualquer desvio externo desconhecido.
Principal Component Analysis (PCA)	Reduz a dimensionalidade de grandes matrizes de logs mapeando as direções de maior variação. Ao reconstruir os pacotes, valores com alto erro de reconstrução indicam anomalias (Rafique *et al*., 2024).	Estratégia: Projeção linear dos dados em eixos ortogonais que maximizam a variância residual. Limitação: Falha ao tentar capturar relações e padrões de ataque estruturados de forma não linear.	Ação: Isola variações anómalas através do cálculo do erro na matriz reconstruída. Foco: Excelente para compressão e limpeza prévia de ruídos de rede.

Autoencoders (AE)	Redes neurais profundas que replicam as entradas na saída passando por um gargalo de compressão. Treinados com tráfego normal, geram erro de reconstrução elevado ao processar ataques (Ahmim *et al*., 2023; Rafique *et al*., 2024).	Estratégia: Aprendizado de representações compactas através de funções de ativação não lineares. Limitação: Complexidade computacional elevada, exigindo hardware dedicado para a borda.	Ação: Dispara alertas quando o erro quadrático de reconstrução ultrapassa o limiar de treino. Foco: Detecção avançada de intrusões complexas e não lineares em IoT.
DBSCAN	Agrupa os dados baseando se na densidade de pontos dentro de um raio específico, dispensando a definição prévia do número de grupos e classificando registros isolados como ataques (Trilles; Hammad; Iskandaryan, 2024).	Estratégia: Expansão de clusters a partir de pontos centrais densos com base em parâmetros de raio. Limitação: Não funciona bem em bases de rede que apresentem densidades muito variáveis.	Ação: Identifica e descarta automaticamente pontos de dados isolados rotulando os como ruído. Foco: Descoberta de agrupamentos de formatos arbitrários na rede.

Agglomerative Clustering	Abordagem hierárquica que constrói uma árvore de agrupamentos de baixo para cima. Os grupos vão se fundindo por critérios de ligação, permitindo rastrear conexões suspeitas fora do escopo usual (Rafique *et al*., 2024).	Estratégia: Fusão sequencial de pares de clusters baseada em matrizes de proximidade. Limitação: Inviável para bases massivas devido à sua complexidade de tempo cúbica.	Ação: Permite uma inspeção detalhada em múltiplos níveis lógicos da hierarquia dos logs. Foco: Análise forense pós incidente de estruturas de tráfego.
Gaussian Mixture Models (GMM)	Assume que os pontos de dados da rede são gerados a partir de uma mistura de distribuições gaussianas, fornecendo uma classificação probabilística refinada para capturar desvios sutis (Rafique *et al*., 2024).	Estratégia: Estimativa de máxima verosimilhança através do algoritmo de Expectação e Maximização. Limitação: Pode convergir para ótimos locais se a inicialização não for bem executada.	Ação: Atribui um grau de probabilidade de pertinência para cada pacote recebido pela rede. Foco: Detecção de desvios de calibração lentos e progressivos em sensores.

Fonte: Elaborado pelos autores (2026).

4.3.2 Métricas de Validação Intrínseca para Models Sem Rótulos

A validação de um classificador inserido no paradigma não supervisionado impõe uma complexidade superior, dada a ausência de uma classe alvo para confrontar os resultados. Na

falta de rótulos para computar matrizes de confusão tradicionais, a avaliação apoia se em critérios de consistência topológica, avaliando as propriedades de coesão interna e separação espacial dos agrupamentos (Trilles; Hammad; Iskandaryan, 2024).

As métricas de validação intrínseca servem como a régua matemática para julgar a estabilidade do modelo (Rafique *et al.*, 2024):

Coeficiente de Silhueta (Silhouette Coefficient)
Avalia a qualidade individual de cada objeto comparando a distância média intracluster com a distância média até o cluster vizinho mais próximo. O cálculo para um elemento é expresso por:

Onde a representa a distância média do ponto para todos os outros elementos do mesmo grupo, e b expressa a menor distância média do ponto para os elementos de qualquer outro cluster. O índice varia de menos um a mais um, onde valores próximos ao limite superior atestam um agrupamento altamente coeso e isolado.

Índice Davies Bouldin (DBI)
Mede a similaridade média entre cada cluster e o seu grupo mais semelhante. A fórmula baseia se na razão entre a dispersão interna dos agrupamentos e a distância geométrica existente entre os centroides desses grupos:

Onde si e sj indicam as dispersões internas dos clusters i e j, enquanto d(ci, cj) representa a distância entre os centroides. Modelos não supervisionados robustos buscam a minimização do índice Davies Bouldin, comprovando grupos distantes e bem definidos.

Índice Calinski Harabasz (Variance Ratio Criterion):
Determina a razão entre a soma dos quadrados entre os clusters e a soma dos quadrados interna dos clusters. A pontuação é computada pela relação matemática:

Onde Tr(B) e Tr(W) correspondem aos traços das matrizes de dispersão entre grupos e dentro dos grupos, N denota o tamanho total da amostra de tráfego e k indica a quantidade de agrupamentos gerados. Valores elevados desse critério validam uma segmentação limpa e estável.

O desbalanceamento massivo de dados afeta sensivelmente esses índices em ambientes de rede (Ahmim *et al*., 2023; Trilles; Hammad; Iskandaryan, 2024). Em fluxos de IoT de plantas automatizadas da Indústria 4.0, o tráfego legítimo pode sufocar geometricamente as amostras isoladas de ataques. Por esse motivo, as investigações contemporâneas usam essas métricas associadas a limiares dinâmicos de distância, garantindo que o modelo não supervisionado preserve a capacidade de isolar anomalias raras sem distorcer as métricas gerais de estabilidade espacial (Rafique *et al*., 2024; Souza; Santos, 2020).

5. CONCLUSÃO

Este estudo realizou uma revisão da literatura focada na análise comparativa entre modelos supervisionados e não supervisionados para a detecção de anomalias em ambientes de Internet das Coisas (IoT). Diante do cenário de alta vulnerabilidade cibernética imposto pelo grande aumento de dispositivos de computação de borda, a implementação de mecanismos baseados em Inteligência Artificial demonstrou ser um pilar indispensável para preservar a integridade e a privacidade das redes contemporâneas.

Os resultados evidenciam que os modelos supervisionados, com destaque para o algoritmo *Random Forest*, exibem patamares surpreendentes de acurácia (frequentemente superiores a 99%) na classificação de ameaças previamente conhecidas e catalogadas. Contudo, seu emprego prático enfrenta uma forte restrição decorrente da escassez e do alto custo de rotulação de dados reais de tráfego, além da ineficácia diante de ataques inesperados. Em contrapartida, as abordagens não supervisionadas, como *Autoencoders* e *Isolation Forest,* consolidam-se como soluções eficazes para mapear o desconhecido, destacando-se pela capacidade de identificar ataques de dia zero sem a necessidade de rótulos prévios.

Além disso, a mudança de arquitetura do processamento centralizado em nuvem para o paradigma do *TinyML* e da computação em névoa (*Fog-Edge Computing*) mostrou-se essencial para viabilizar a segurança em tempo real. A execução local de inferências diretamente em microcontroladores de recursos restritos mitiga drasticamente os problemas de atraso e uso de rede, tornando a defesa mais rápida, proativa e distribuída. Notou-se, contudo, que o cenário atual da pesquisa avança para a combinação de modelos, unindo características de diferentes redes profundas para alcançar alta performance.

Por fim, conclui-se que não existe um algoritmo universal; a seleção do modelo ideal permanece estritamente condicionada à topologia física do ecossistema, à criticidade temporal da aplicação e às restrições de hardware dos dispositivos embarcados.

REFERÊNCIAS

AL-QUAYED, F. AI-Powered Anomaly Detection and Cybersecurity in Healthcare IoT with Fog-Edge. **Computer Modeling in Engineering & Sciences**, [*S. l.*], v. 146, n. 1, p. 1–10, 2026. Disponível em: <https://www.techscience.com/CMES/v146n1/65732>. Acesso em: 14 jun. 2026.

ZENG, H. *et al.* Towards a conceptual framework for AI-driven anomaly detection in smart city IoT networks for enhanced cybersecurity. **Journal of Innovation & Knowledge**, [*S. l.*], v. 9, n. 4, p. 100601, 1 out. 2024. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2444569X24001409>. Acesso em: 14 jun. 2026.

AHMIM, A. *et al.* Distributed Denial of Service Attack Detection for the Internet of Things Using Hybrid Deep Learning Model. **IEEE Access**, [*S. l.*], v. 11, p. 119862–119875, 1 jan. 2023. Disponível em: <https://ieeexplore.ieee.org/document/10295488>. Acesso em: 14 jun. 2026.

TRILLES, S.; HAMMAD, S. S.; ISKANDARYAN, D. Anomaly detection based on Artificial Intelligence of Things: A Systematic Literature Mapping. **Internet of Things**, [*S. l.*], v. 25, p. 101063, 1 jan. 2024. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S2542660524000052>. Acesso em: 14 jun. 2026.

RAFIQUE, S. H. *et al.* Machine Learning and Deep Learning Techniques for Internet of Things Network Anomaly Detection—Current Research Trends. **Sensors**, [*S. l.*], v. 24, n. 6, p. 1968, 1 jan. 2024. Disponível em: <https://www.mdpi.com/1424-8220/24/6/1968>. Acesso em: 14 jun. 2026.

SOUZA, Marina Teixeira de; SANTOS, Fernando César Almada. Competências Operacionais e Industria 4.0: revisão sistemática da literatura. Future Studies Research Journal: Trends and Strategies, v. 12, n. 2, p. 264-288, 1 maio 2020. Disponível em: <https://futurejournal.org/FSRJ/article/view/499>. Acesso em: 14 jun. 2026.

MAIA, Rebeca; LEMOS, Sandra. Usando Ciência de Dados para detectar anomalias em logs de sistema. **Insight Data Science Lab**, Fortaleza, 21 jun. 2021. Disponível em: <https://www.insightlab.ufc.br/usando-ciencia-de-dados-para-detectar-anomalias-em-logs-sistema/>. Acesso em: 14 jun. 2026.