FACULDADE DE TECNOLOGIA DE SÃO PAULO

GABRIEL ZANGERME

APLICAÇÕES DE MACHINE LEARNING EM CYBER THREAT INTELLIGENCE: UMA ANÁLISE SISTEMÁTICA DE BIBLIOGRÁFIAS

SÃO PAULO

2022

FACULDADE DE TECNOLOGIA DE SÃO PAULO

GABRIEL ZANGERME

APLICAÇÕES DE MACHINE LEARNING EM CYBER THREAT INTELLIGENCE: UMA ANÁLISE SISTEMÁTICA DE BIBLIOGRÁFIAS

Trabalho submetido como exigência parcial
para a obtenção do Grau de Tecnólogo em
Análise e Desenvolvimento de Sistemas
Orientador: Dr. Carlos Hideo Arima

SÃO PAULO

2022

FACULDADE DE TECNOLOGIA DE SÃO PAULO

GABRIEL ZANGERME

APLICAÇÕES DE MACHINE LEARNING EM CYBER THREAT INTELLIGENCE: UMA ANÁLISE SISTEMÁTICA DE BIBLIOGRÁFIAS

Trabalho submetido como exigência parcial para a obtenção do Grau de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Parecer do Professor Orientador:_____

_____

_____

Orientador: Dr. Carlos Hideo Arima

SÃO PAULO, 12 de dezembro de 2022.

"The purpose of life is finding the largest burden that you can bear and bearing it"
(Jordan Peterson)

RESUMO

Cyber Threat Intelligence trata-se do processo de obtenção de conhecimento acionável sobre atores de ameaças cibernéticas e atividades maliciosas, possibilitando com que através do conhecimento obtido, organizações possam complementar os componentes de sua estrutura de cyber security, auxiliando em todas as etapas do processo, podendo reduzir os danos ocasionados por ameaças cibernéticas.

Tratando-se de um processo que envolve a tomada de decisão, técnicas de machine learning podem auxiliar em Cyber Threat Intelligence, possibilitando a coleta de insights e proporcionando uma operacionalização de suas etapas de forma mais eficiente, dado sua utilização para detecção e classificação de ameaças de forma mais rápida e baseada em dados.

Este trabalho realiza uma análise textual, baseando-se em um livro considerado uma referência no assunto de técnicas de machine learning para Cyber Security, seguido de uma análise bibliométrica utilizando as palavras com maior frequência na obra escolhida como referência.

Baseando-se nos artigos encontrados, é realizada uma revisão sistemática sobre os temas abordados, para identificar como o tema de machine learning é abordado no domínio de Cyber Security, quais as técnicas são mais utilizadas e como são aplicadas.

Após a análise dos artigos coletados, foi possível identificar os modelos mais utilizados e suas aplicações, e afirmar a proposição de que o Random Forest (RF) é o modelo predominante na literatura dentre os outros modelos levantados.

O Random Forest apresenta uma grande utilização devido a sua versatilidade de aplicação em casos de detecção de classificação de ameaças, além de possuir uma fácil implementação, conseguindo um desempenho superior até mesmo a técnicas mais avançadas e robustas de machine learning, como observado na detecção de Malwares, onde teve uma acurácia de 99.78% comparado a 99.21% em a relação a implementações semelhantes de um Deep Neural Networks (DNNs).

Com base nas aplicações de machine learning no domínio de Cyber Security que podem auxiliar no processo de Cyber Threat Intelligence, nota-se uma grande variedade de nichos e tecnologias que podem se beneficiar da implementação destes algoritmos, garantindo um maior entendimento do comportamento de suas ameaças, e aumentem sua resiliência contra ataques cibernéticos.

Palavras Chave: Cyber Threat Intelligence, Threat Intelligence, Machine Learning, Random Forest

ABSTRACT

Cyber Threat Intelligence is the process of gaining actionable knowledge about cyberthreat threats and malicious activities, enabling organizations to complement the components of their cybersecurity framework through the knowledge gained, assisting at all stages of the process, which can reduce the damage caused by cyber threats.

As it is a process that involves decision-making, machine learning techniques can assist in Cyber Threat Intelligence, allowing the collection of insights and providing a more efficient operationalization of its steps, given its use for detection and faster data-driven threat classification.

This work performs a textual analysis, based on a book considered a reference in the subject of machine learning techniques for Cyber Security, followed by a bibliometric analysis using the most frequent words in the work chosen as a reference.

Based on the articles found, a systematic review is carried out on the topics covered, to identify how the machine learning topic is likely to be in the field of Cybersecurity, which techniques are most used and how they are applied.

After analyzing the collected articles, it was possible to identify the most used models and their applications, and to affirm the proposition that the Random Forest (RF) is the predominant model in the literature among the other models surveyed.

Random Forest is widely used due to its application versatility in cases of threat classification detection, in addition to being easy to implement, it supported superior performance even for more advanced and robust machine learning techniques, as observed in the detection of machine learning Malwares, where it had an accuracy of 99.78% against 99.21% in relation to similar implementations of Deep Neural Networks (DNNs).

Based on machine learning applications in the Cyber Security domain that can help in the Cyber Threat Intelligence process, there is a wide variety of niches and technologies that can benefit from the implementation of these algorithms, ensuring a greater understanding of the behavior of their threats, and increasing the resilience against cyberattacks.

Keywords: Cyber Threat Intelligence, Threat Intelligence, Machine Learning, Random Forest

SUMÁRIO

1. INTRODUÇÃO

Cybersecurity, trata-se da disciplina de proteger redes, dispositivos e dados contra acessos ilícitos, ou uso criminoso, garantindo confidencialidade, integridade e disponibilidade das informações (CISA, 2021).

Empresas, e organizações, assim como órgãos de vários governos são prejudicados pelo roubo de propriedade intelectual, segredos comerciais e outras informações que possam ser vitais a uma instituição ou altamente valiosas (CISA, 2021).

Caso pudesse ser medido como um país, os danos causados por crimes cibernéticos, que em em 2021 excedeu a valor de 6 trilhões de dólares, poderia ser considerado a terceira maior economia do mundo, atrás somente dos Estados Unidos e China. Tendo expectativas deste número, aumentar em 15% até 2025 (CYBERCRIME MAGAZINE, 2020).

Cyber Threat Intelligence, trata-se do conhecimento agregado, proveniente da coleta, análise e processamento de dados de Cybersecurity, garantindo um melhor entendimento do comportamentos de agentes de ameaça, alvos de ataques e motivações, permitindo que organizações e suas equipes de segurança, possam adotar uma postura proativa em relação a disciplina de Cybersecurity, tomando decisões baseadas em dados e garantindo maior assertividade ao lidar com ataques cibernéticos (SPLUNK, 2022).

Junto ao panorama de um crescente número de crimes cibernéticos e de esforços para melhor entender cada agente de ameaça, a disciplina de Inteligência Artificial e seus subdomínios, como Machine Learning, desempenha um papel fundamental, fazendo com que o processo de Cyber Threat Intelligence possa trabalhar de forma mais efetiva com grandes volumes de dados (OUTPOST24, 2020) e garantindo um melhor desempenho do trabalho conjunto entre humano e máquina, possibilitando uma visão mais profunda do que realmente possa ter acontecido durante um incidente de Cyber Security, ao inves e seguir processos de triagem manual, evitando em muitos casos, a necessidade de executar outras ferramentas forenses. (SENTINELONE, 2022).

Com a crescente utilização da disciplina de Inteligência Artificial em Cyber Security, cada vez mais organizações se movem para refinar técnicas que possam colocá-las em uma posição favorável diante a ameaças cibernéticas e seus impactos.

Neste cenário, quais técnicas de machine learning são mais utilizadas na identificação dos métodos, vetores e técnicas utilizadas por atores de ameaça, garantindo um processo de Cyber Threat Intelligence com melhor tempo de resposta e maior assertividade para contextualizar qual o tipo de ameaça, aplicando-se assim um plano de resposta adequado?

Este trabalho realiza uma análise textual, baseando-se em um livro considerado uma referência no assunto de técnicas de machine learning para Cyber Security, seguido de uma análise bibliométrica utilizando as palavras com maior frequência na obra escolhida como referência.

Baseando-se nos artigos encontrados, é realizada uma revisão sistemática sobre os temas abordados, para

identificar como o tema de machine learning é abordado no domínio de Cyber Security, quais as técnicas são mais utilizadas e como são aplicadas.

Como contribuição, este trabalho tem como propósito, identificar os modelos e técnicas de machine learning mais utilizados no domínio de Cyber Security e que possam dar suporte a processos de tomada de decisão em Cyber Threat Intelligence, assim como suas aplicações, com a proposição de que, pelo fato de ser um modelo de fácil implementação, alto desempenho, tanto para classificação quanto categorização, além de ser amplamente utilizado em outros setores como financeiro, saúde e fraudes, o Random Forest pode ser o modelo predominante dentro da literatura entre as outras técnicas existentes.

A identificação de modelos e técnicas de machine learning que são mais comuns em Cyber Security, e como estas mesmas são aplicadas, auxiliam no entendimento de que quais modelos são mais apropriados para cada cenário, proporcionando tomadas de decisão mais assertivas no processo de Cyber Threat Intelligence, aprimorando as decisões antecipatórias e até a reação aos atores maliciosos, optimização de operacionalizações, e até mesmo entendimento proativo de futuras ameaças. Tais fatores contribuem para que organizações e empresas aumentem sua resiliência contra ataques cibernéticos, permitindo também a otimização de seus esforços e custos para a construção de um ambiente mais seguro.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 CYBER THREAT INTELLIGENCE

Cyber Threat Intelligence tem se tornado cada vez mais um tópico relevante quando discutido no domínio de Cyber Security. Devido à falta de literaturas mais concisas sobre o tema, o termo muitas vezes é adaptado para definições específicas baseadas em diferentes pontos de vista procedurais e comerciais (NATIONAL CYBER SECURITY CENTER, 2015).

(CALTAGIRONE, 2018), define a Cyber Threat Intelligence (também citado como Threat Intelligence), como o conhecimento acionável sobre agentes adversários e suas atividades maliciosas, possibilitando com que organizações reduzem os danos ocasionados por estes mesmo a partir de uma tomada de decisão mais segura.

(PLANQUE, 2017), cita a definição de Michael Cloppert, a qual propõe uma definição baseada em três partes sobre o que é Cyber Threat Intelligence.

1. "Eu defino operações de Cyber Threat Intelligence, como ações tomadas no domínio de Cybersecurity, para comprometer e defender informações e recursos protegidos disponíveis naquele domínio."
2. "Eu defino Análise de Cyber Threat Intelligence como a análise destas ações e o seus atores, ferramentas e técnicas por trás para apoio das operações."
3. "Eu defino o domínio de Cyber Threat Intelligence como a união das Operações e Análise de Cyber Threat Intelligence"

Em contraste, (R.M. LEE, 2016) define Cyber Threat Intelligence como “O processo e o produto resultante da interpretação de dados brutos em informações que atendem a um requisito, no que se refere aos atores que têm a intenção, oportunidade e capacidade de fazer mal”.

Podemos analisar que as duas definições se complementam, com a proposta de Cloppert englobando todos os atributos intrínsecos do processo, ferramentas e ações necessárias para executar a disciplina e a

proposta de Lee abrangendo o contexto do uso de dados para se prevenir no que se refere a ameaças cibernéticas. Cloppert também engloba que a disciplina de Cyber Threat Intelligence propicia não só a defesa, mas também explora a possibilidade do uso de inteligência para uma postura mais ofensiva, tratando-se do domínio de Cyber Security.
O processo de Cyber Threat Intelligence, permite com que, integrado a um sistema de segurança, reduza o tempo médio de recuperação durante um ataque cibernético, e também, um maior preparo maior para que organizações consigam se proteger, entendendo como agentes maliciosos interrompem ou comprometem sistemas e antecipando-se em relação aos mesmos de forma proativa, durante e depois de um incidente (CALTAGIRONE, 2018).

Para responder a um incidente, deve haver um método para que se possa adquirir conhecimento sobre a ameaça, sendo que todo processo de Cyber Threat Intelligence deve responder a três perguntas principais:

1. Qual é a ameaça? Dando embasamento para que endereçar entendendo quando, como e porque se deu a ameaça.
2. Qual será o impacto para a organização, caso a ameaça se concretize?
3. Quais ações podem mitigar a ameaça em questão a curto e médio prazo?

Cyber Threat Intelligence consegue responder a essas perguntas a partir da definição do contexto da ameaça, preocupando-se principalmente com o porquê da mesma estar ocorrendo e qual ação deve ser tomada. O processo de definição de contexto sem ações, faz com que o processo de Cyber Threat Intelligence seja vago e traga pouco valor com os dados coletados.

No contexto de Cyber Threat Intelligence, ações, tratam-se de recomendações técnicas e políticas de segurança específicas para cada tipo de ameaça, avaliando seu impacto e comportamento. Este processo abrange desde detalhes técnicos que permitem a detecção destas ameaças, como demais insights estratégicos que possam vir a ser úteis para tomada de decisão de executivos de uma corporação.

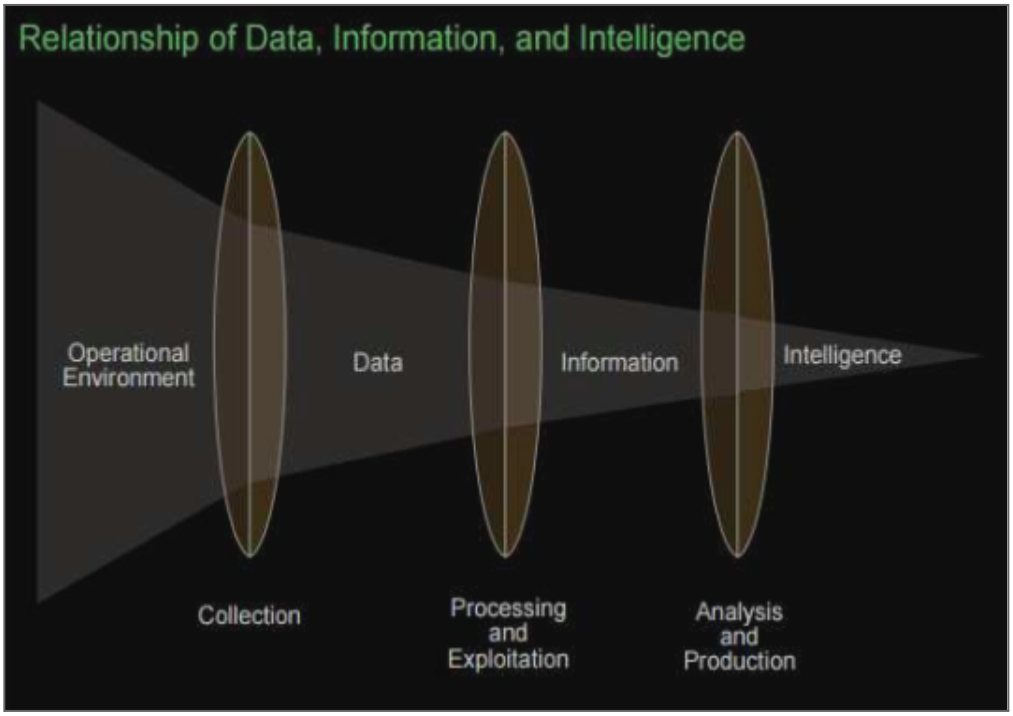
Entre os processos de ação, destacam-se dois:

1. Indicators of Compromise (IOCs): Composto por elementos técnicos de informação, como protocolos de IP, nomes de domínio, nomes de arquivos e hashes, utilizados principalmente para detecção da ameaça, seja de maneira proativa ou reativa.
2. Threat Behavior Analytics: Tem como objetivo identificar interações de sistemas ou usuários que possam ser suspeitas. Este tipo de abordagem traça o comportamento ao invés de aspectos puramente técnicos como no IOCs, isso traz uma vantagem em relação a atores de ameaça, pois comportamentos maliciosos, diferente de traços técnicos, são mais difíceis de se modificar, demandando assim um tempo maior para adaptação de novos tipos de ataques e ameaças.

O processo de Cyber Threat Intelligence sozinho, não pode proteger uma corporação de ameaças como um todo, mas a sua utilização, complementa todos os componentes de uma estrutura de cyber security, auxiliando em todos os estágios: detecção, prevenção e resposta. Quando usada apropriadamente, pode reduzir os danos a uma organização, justamente por aumentar a eficácia de uma organização (CALTAGIRONE, 2018).

Para construção de um processo ou produto de inteligência, é necessário haver um modelo por meio de identificação, coleta e análise de informações e dados (CISA, 2020). Na Figura 1 é possível observar todo o ciclo o qual é percorrido até chegar na inteligência de um processo.

Figura 1 - Relação Entre Dados, Informação e Inteligência.



Fonte: (USA JOINT INTELLIGENCE, 2013)

Diversos processos de Inteligência baseiam-se nesse mesmo ciclo, adaptando para contextos de cada organização e apresentando resultados muito efetivos, tratando-se do domínio de Cyber Security. O CISA (Cybersecurity & Infrastructure Security Agency), propõe um framework de Cyber Threat Intelligence baseado no ciclo de inteligência padrão, mas adicionando um enfoque maior na intersecção do domínio de Cyber Security e tecnologias. Esse mesmo framework (CISA, 2020), é constituído por:

1. Contexto de Ambiente: Compreensão da organização, quais são suas ameaças e riscos que podem ser exploradas dada a natureza de sua indústria e como são constituídas as operações internas e externas da sua organização.
2. Obtenção de Dados: Através de automações, os dados e informações são coletadas de várias fontes internas e externas para analistas realizarem análises e responderem a requisitos de inteligência organizacional.
3. Análise de Ameaças: Abrange tanto análise tática quanto análise operacional, levando em conta ameaças específicas, ataques, incidentes, vulnerabilidades e contextos gerais de ameaças pertinentes a alguma indústria ou nicho, que possam apoiar na tomada de decisão proativa e/ou responsiva.
4. Análise Estratégica: Processo de análise holística sobre as ameaças, levando em conta seu potencial de risco, nível de exposição da organização que está sofrendo o ataque e o impacto organizacional caso esse tipo de ameaça se concretize.
5. Relatórios e Feedback: Estabelecer a partir de todo o framework, uma comunicação clara entre os analistas envolvidos no ciclo e os responsáveis pela tomada de decisão.

O framework proposto pela CISA, além dos passos comumente implicados em um processo de Cyber Threat Intelligence, que utiliza-se do acumem analitco humano, adiciona a interação humano e computador, sendo esta última, a utilização de modelos de machine learning supervisionados ou não supervisionados.

A rapidez, poder computacional e possibilidade de melhoria contínua baseada em dados, que é observada na utilização de modelos de machine learning, potencializam o processo de inteligência, gerando uma maior eficiência em todas as etapas no processo, auxiliando na detecção, categorização e predição de ameaças e também podendo auxiliar na etapa de relatórios e Feedbacks.

Como exemplo dos benefícios da intersecção dessas duas disciplinas, em outro experimento propondo um framework de Cyber Threat Intelligence para FinTechs (NOOR et al, 2019), observou que a utilização de modelos de machine learning conseguiram prever perfis de CTAs (Cyber Threat Agents) com uma precisão de 83% comparado a 33% dos sistemas baseados em regras.

2.2 MACHINE LEARNING

Machine Learning é um subcampo da Inteligência Artificial (AI), definido como a capacidade de uma máquina imitar o comportamento humano inteligente. O processo para construção de um modelo, como também é chamado, começa com a coleta e preparo dos dados, que podem ser de diferentes fontes e formatos, o dado então é utilizado para o treino do modelo, sendo que quanto mais dados de exemplo para o modelo, melhor será sua performance.

A função de um modelo de machine learning pode ser descritivo, ou seja, ele usa os dados para explicar o que já aconteceu; preditivo, utilizando-se dos dados para predizer o que irá acontecer no futuro ou prescritivo, o que significa que o modelo irá fazer sugestões sobre qual ação deve ser tomada (MIT MANAGEMENT SLOAN SCHOOL, 2021).

Os modelos de machine learning podem ser divididos em 3 grupos, baseados no tipo de aprendizado que estes modelos exercem no processo de treinamento a partir dos dados (MIT MANAGEMENT SLOAN SCHOOL, 2021):

1. Aprendizado Supervisionado: modelos que são treinados com dados rotulados, o que proporcionam com que o modelo aprenda e tornem-se mais precisos com o tempo.
2. Aprendizado Não Supervisionado: modelos que procuram por padrões em dados não rotulados, sendo assim, este tipo de algoritmo tenta procurar por tendências que não foram explicitadas pelas pessoas em sua criação.
3. Aprendizado por Reforço: modelos que aprendem por tentativa e erro, pegando assim parâmetros sobre qual é o melhor caminho a se tomar por conta própria, muito parecido com o processo de aprendizagem humano.

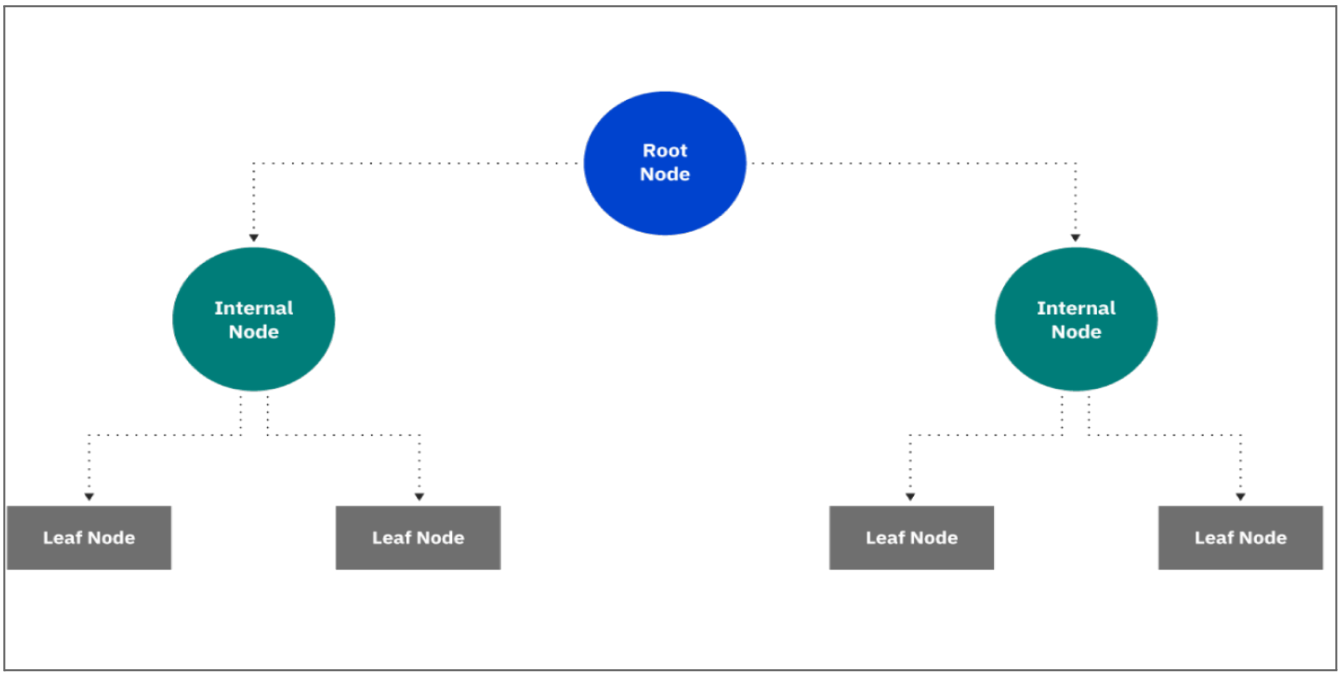
2.2.1. DECISION TREES

Decision Trees são um algoritmo de machine learning de aprendizado supervisionado, podendo ser

implementado em problemas de regressão e classificação (KDNuggets, 2022).

Este tipo de algoritmo segue um formato hierárquico de árvore, o qual consiste em um nodo principal (Root Node), ramos internos (Branches), nós internos (Internal Node) e nós folha (Leaf Node).
A Figura 2 ilustra a hierarquia a qual consiste o funcionamento das Decision Trees.

Figura 2 - Hierarquia de Árvore



Fonte: (IBM, 2020)

Como pode-se observar na Figura 2, o algoritmo de Decision Tree começa a partir do Root Node, que não possui nenhuma ramificação de entrada. A partir do Root Node, os Internal Nodes, também conhecidos como Deivison Nodes (Nós de Decisão), com base em características disponíveis, realiza avaliações para formar subconjuntos homogêneos, observados na Figura 2 como Leaf Nodes. Os Leaf Nodes representam todos os resultados possíveis dentro do conjunto de dados imputados a partir do Root Node (IBM, 2020).

Decision Trees empregam uma estratégia em seu algoritmo de dividir e conquistar, conduzindo assim uma busca com grande insumo para identificar os pontos de divisão ideal dentro da estrutura de árvore. Este processo é repetido de maneira descendente recursiva até que a maioria dos registros imputados tenham sido classificados.

O algoritmo de Decision Trees apresenta como vantagem a sua fácil interpretação e implementação, assim como uma grande flexibilidade dado a sua característica para problemas de regressão e classificação. No enteado, à medida que uma árvore cresce em tamanho na sua aplicação, torna-se cada vez mais difícil manter a pureza de resultados, resultando em poucos dados caindo em uma determinada subárvore, ocorrendo assim o fenômeno de fragmentação de dados, o que pode ocasionar em overfitting. (IBM, 2020.

O overfitting trata-se de uma disfunção em modelos de machine learning, na qual o modelo fornece previsões ou categorizações precisas, porém somente para os dados utilizados no treino do modelo, não oferecendo a mesma precisão para novos dados (AWS, 2022).

O Decision Trees mostra-se bastante eficiente para detecção de ameaças, em cenários de aplicabilidade de detecção de invasões, o algoritmo junto a softwares de fácil implementação de modelos de machine learning como o Weka, conseguiu atingir uma acurácia de 98% para identificar invasões e atividades maliciosas em ambientes de rede (MARKEY, 2011).

2.2.2. RANDOM FOREST

Dentre os modelos de machine learning, o modelo de Random Forest (RF) destaca-se como um algoritmo supervisionado robusto, amplamente utilizado para problemas de regressão e categorização (ANALYTICS VIDHYA, 2021).

O algoritmo de Random Forest é composto por vários conjuntos de classificadores, no caso, Decision Trees, e suas previsões são agregadas para identificação do resultado mais presente. Em 1996, Leo Breiman, apresentou um novo método, conhecido como Bagging, onde uma amostra de dados aleatórios em um conjunto de dados utilizado para treino é selecionada com sobreposição, o que significa que os pontos de dados podem ser escolhidos mais de uma vez. Após várias amostras de dados serem geradas, esses modelos de Decision Trees são treinados de forma independente, e dependendo de qual seja o propósito de aplicação, ou seja, regressão ou classificação, a média ou a maioria das previsões de classificações, produzem uma estimativa. Este tipo de abordagem, é comumente usada para reduzir a variância em um conjunto de dados que apresentem informações irrelevantes, ou dados nulos, também conhecidos como "*noisy data"* (IBM, 2020). Ao levar em conta toda a variabilidade potencial dos dados, pode-se reduzir o risco de overfitting, viés e variância geral, resultando em previsões mais precisas.

O Random Forest, é uma extensão ao método de Bagging, pois baseia-se na estrutura do mesmo, com a adição do método de Feature Randomness. Feature randomness, também conhecido como "random subspace method", gera um subconjunto aleatório de features, o que garante uma baixa correlação entre as Decision Trees do método de Bagging. Esse ponto, implica um diferencial fundamental entre Random Forest e métodos comuns de árvore de decisão. Enquanto Decision Trees consideram todas as divisões de features possíveis, inclusive aqueles que possam impactar na performance do modelo, o Random Forest seleciona apenas um subconjunto dessas features, garantindo um modelo mais preciso e menos enviesado (IBM, 2020).

Random Forest também possui uma alta capacidade de detecção de categorização em ambientes que possam haver sobreposição de características, como exemplo ambientes que possuam muitos veículos e é necessário categorizar apenas uma marca específica, ou principalmente na identificação de ameaças cibernéticas, onde vários ataques podem dividir características de comportamento (NOOR et al, 2019).

O Random Forest, apesar de ser um algoritmo robusto e que apresenta bons resultados e aplicabilidade, necessita de um poder computacional menor, comparado a outros modelos avançados que apesar de apresentarem resultados também excepcionais, apresentam um custo benefício menor, dado a necessidade de um volume de dados e poder computacional maior, como exemplo a aplicação de Deep Learning (SEWAK;SAHAY;RATHORE,2018).

Em experimentos de detecção de anomalias relacionadas a ataques em ambientes multi-cloud, o Random Forest apresentou uma detecção de ataques próxima a 99% e uma acurácia próxima a 94% para a classificação do tipo de ameaça (SALMAN et al.,2017) pontos que podem auxiliar no processo de tomada de decisão em um framework de Cyber Threat Intelligence e possibilitando assim uma ação reativa mais rápida e abrindo a possibilidade para processo proativos de interceptação de ameaças.

Em frameworks de Cyber Threat Intelligence voltados para detecção de URLs maliciosos, a utilização do Random Forest, trouxe um aumento significativo de 7.8% na acurácia e 6.7% de redução de falsos positivos, comparado a métodos convencionais de detecção já existentes (ALSAEDI et al, 2022).

2.2.3. K-NEAREST NEIGHBORS

O K-Nearest Neighbors (KNN) está entre os algoritmos de machine learning de aprendizagem supervisionada mais simples, sendo amplamente estudado no campo de reconhecimento de padrões ao longo do último século. Embora não seja tão popular quantos outros algoritmos usados em alta escala, ele ainda é amplamente usado, sendo recomendado principalmente em problemas que envolvam classificação como um benchmark de desempenho preditivo no desenvolvimento de modelos mais sofisticados (University of Wisconsin, 2018).

O KNN utiliza-se de estimativas para fazer classificações ou regressões sobre o agrupamento de pontos de dados individuais utilizados. Sua utilização mais comum em problemas de classificação parte do pressuposto de que pontos de dados semelhantes podem ser encontrados próximos uns dos outros (IBM, 2020).

Para classificação, um rótulo de classe é atribuído com base no conceito de "*majority vote"*, ou seja, o rótulo que é representado com maior frequência em torno de um determinado ponto de dados é utilizado. Embora esse termo possa ser considerado como algo chamado de *"plurality voting",* o termo "majority vote" é mais observado na literatura. As terminologias apesar de semelhantes, distinguem-se, pois o "majority vote" implica que uma maioria de classificações sobre um rótulo esteja acima de 50%, o que é funcional principalmente em casos de categorização de apenas duas categorias. Quando várias classes são encaixadas em um problema, envolvendo assim mais de duas categorias, não existe a necessidade de 50% de votos sobre um grupo de dados para que haja uma conclusão sobre sua classificação, sendo em um caso de quatro categorias, a necessidade de apenas 25% dos votos dos classificadores (IBM, 2020). A Figura 3 apresenta de forma resumida o funcionamento do KNN para as classificações.

Figura 3 - Funcionamento de Classificação do KNN

Fonte: (University of Wisconsin, 2018)

Na Figura 3 é possível observar o processo de classificação do K-Nearest Neighbours. O ponto roxo, representando um novo dado de entrada é então comparado com as classes A e B próximas, representadas pelos pontos Azul e Verde respectivamente. O processo de *"Majority Vote"* mostra que o novo dado de entrada está mais próximo de classes B do que classes A (Duas classes B contra Uma classe A), sendo assim o novo dado de entrada roxo, é classificado como classe B tornando-se um novo ponto verde.

O KNN tem como vantagem a sua fácil implementação, necessita de menos parâmetros comparado a outros algoritmos de machine learning, nesse caso, é apenas utilizado a distância que será considerado entre as classes como parâmetro principal, e sua versatilidade, dado que apesar de ser majoritariamente utilizado para classificações, pode também ser usado para problemas de regressão. Em contrapartida o KNN sofre para cenários que exigem a necessidade de escalar para um número maior de dados, devido a sua característica de ser um Lazy Algorithm, ou seja, o processo de aprendizado é generalizado, não sendo executado anteriormente levando como base o dataset de treino, aplicando-o somente quando uma nova requisição de predição ou classificação é executada, o que implica em um tempo de resposta maior e maior poder computacional (IBM, 2020).

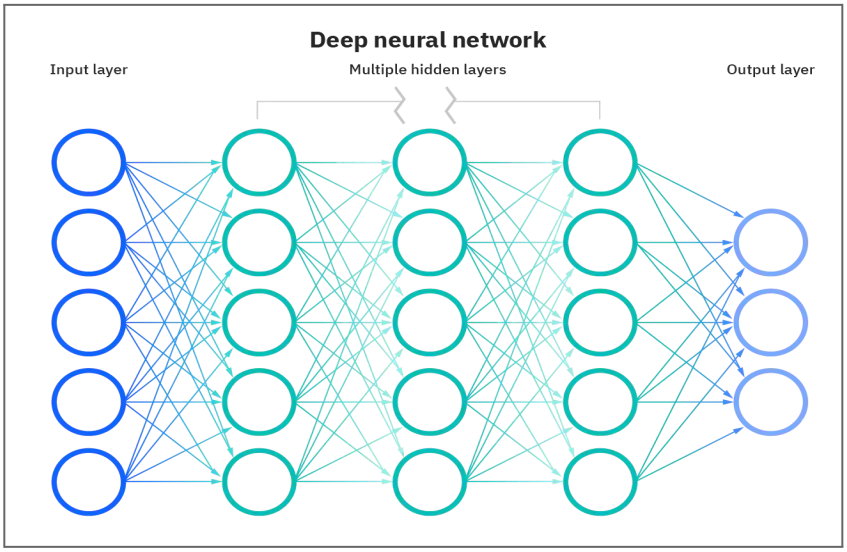
Em aplicações relacionadas a detecção de invasões, o KNN mostra-se eficiente quanto a classificação em ambientes dinâmicos os quais requerem atualizações frequentes no dataset de treino do modelo, o que pode ser atrativo quando se olha o cenário de detecção de invasões (LIAO; VEMURI, 2002).

2.2.4. DEEP NEURAL NETWORKS

O Deep Neural Networks (DNNs) é uma abordagem na qual pertence um subgrupo dos algoritmos clássicos de machine learning, comumente referenciada como Deep Learning, sendo essencialmente um Neural Network com três ou mais layers.

Layers são nós que servem para representar camadas de entrada, processamento e saída de dados, como pode ser observado na Figura 4.

Figura 4 - Funcionamento de Layers de Deep Neural Networks



Fonte: (IBM, 2020)

Na Figura 4, os nós de entrada de dados (Input layer) são apresentados como o primeiro layer, nele são inseridos os dados que posteriormente serão utilizados em nós de processamento (Hidden layers), sendo estes, os responsáveis pelas regras de treino do algoritmo, podendo ser construído um ou mais hidden layers dependendo do processamento. O último layer (Output layer), representa a camada de saída do processamento, onde de fato é passado o resultado final de uma regressão ou categorização.

Esse tipo específico de Neural Network tenta simular o comportamento do cérebro humano, permitindo com que esse tipo de abordagem possa aprender a partir de grandes volumes de dados (IBM, 2020)

As DNNs diferem-se dos demais algoritmos de machine learning devido a como esse tipo de algoritmo aprende com os dados que lhe são apresentados. Enquanto algoritmos clássicos lidam com dados estruturados e dados rotulados para fazer predições ou classificações, o que implica na pré-definição de features quando os dados são imputados, os DNNs podem ingerir e processar dados não estruturados como textos, imagens e automatizar o processo de extração de features, removendo algumas dependências e implementações no processo tradicional de modelagem de algoritmos de machine learning. (IBM, 2020).

Devido a grande complexidade deste tipo de algoritmo, o mesmo possui variações e adaptações em suas estruturas para atender a cenários específicos, como no caso das Convolutional neural networks (CNNs), utilizadas primariamente para classificação de imagens e visão computacional e as Recurrent Neural Networks (RNNs), utilizadas principalmente para o processamento de linguagem natural e reconhecimento de voz (IBM, 2020).

O grande poder de abstração das DNNs e processo intrínseco de seleção de features, traz consigo desvantagens, sendo estas a necessidade de volumes grandes de dados para treino, e a necessidade de alto poder computacional o que implica automaticamente no custo para implantar este tipo de modelo em cenários produtivos (SMEBOOK, 2021).

Deep Neural Networks estão amplamente aplicadas no domínio de Cyber Security, auxiliando na detecção de padrões de invasões em redes, detecção de malwares para auxiliar na melhoria de firewalls e até mesmo prevenindo spams que possam conter ataques que envolvam engenharia social via emails e SMS (DATTO, 2022).

2.2.5. MÉTRICAS DE AVALIAÇÃO DE MODELOS

Além da utilização apropriada de modelos diferentes para cada tipo de cenário, e problema a ser abordado, um passo importante na implementação de machine learning, é a avaliação de métricas, para possibilitar a comparação de desempenho de um modelo em treinamento e/ou semelhantes.

Entre as principais métricas que podem ser avaliadas, observa-se como principais (KD):

- Acurácia: Representa o total de acertos de um modelo (Verdadeiro Positivos e Verdadeiros Negativos), sobre o número total de tentativas de predição ou classificação.
- F-Score: Calculo como a média harmônica entre a precisão e o recall, geralmente utilizado em casos onde há uma aproximação de ambas as métricas, dificultando a avaliação das métricas.
- Recall: Representa a sensibilidade de classificação de um modelo, trazendo a porcentagem de acerto de Verdadeiros Positivos em relação ao número de classificações (Verdadeiros Positivos e Falsos Negativos).
- Precisão: Utilizado para trazer a porcentagem de acerto de Verdadeiros Positivos em relação ao número total de classificações (Verdadeiros Positivos e Falsos Positivos.

3. METODOLOGIA

A fim de estabelecer a bibliometria com consistência, foi escolhido um referencial teórico, que segundo (MARCONI; LAKATOS, 2003), auxilia a identificar o estado do problema que será abordado na pesquisa, levando em consideração estudos e pesquisas já pré estabelecidas.

Tratando-se da disciplina de Machine Learning no domínio de Cyber Security, o livro de Sumeet Dua e Xian Du, *"Data Mining and Machine Learning in Cybersecurity"*, de 2016, pode ser considerado como uma referência para o tema. A obra possuía um total de 512 citações, na data de 01/09/2022, data a qual a bibliometria foi realizada.

Para a análise textual, foi utilizado um código em Python na versão 3.9, junto a outras bibliotecas para facilitar em cada uma das etapas, sendo estas: PdfMiner para extração de texto em arquivos PDF, Textblob para simplificar estruturas textuais, Numpy e Pandas para criação de Arrays e estruturas de leitura, Nltk para remoção de stopwords e palavras indesejadas, e o Matplotlib e WordCloud para plotagem de gráficos e nuvem de palavras.

Com o código e bibliotecas apresentadas, foram processados os textos coletados a partir do livro de referência, gerando uma nuvem de palavras, como pode-se observar na Figura 5, para facilitar a identificação de palavras mais citadas na obra escolhida como referencial teórico. Com base no resultado, foi possível realizar a bibliometria sobre o assunto.

Figura 5 - Nuvem de Palavras

Fonte: Resultado da Pesquisa

Ao observar a nuvem de palavras, é possível verificar que as palavras mais frequentes no referencial teórico e que se encaixam na proposta desta pesquisa foram: *Machine Learning, Learning Method e Intrusion Detection*, neste caso, são desconsiderados termos mais genéricos como: *data, set e network traffic, data mining,* que não se enquadram na proposta deste trabalho.

Com as a captura das palavras mais frequentes e que se encaixam na proposta de pesquisa, foi realizada a bibliometria sobre o tema. A mineração dos dados foi realizada utilizando o software *"Publish or Perish v8"*, nada data de 05/09/22, com os seguintes parâmetros apresentados na sequência:

Como título, as palavras Cyber Threat Intelligence foram atribuídas. Como palavras chave, foram utilizadas as palavras *Machine Learning, Learning Method e Intrusion Detection*. Utilizando-se *and* entre as palavras chave.

Os Databases utilizados para mineração dos dados no "*Publish or Perish"* foram: Google Scholar, Crossref e Scopus.

Foram então incluídos os registros publicados entre 2016 e 2022, disponíveis no idioma Inglês e Português. Totalizando 1225 resultados.

Com a mineração dos dados utilizando as palavras chaves, período e linguagens desejadas, após os resultados, foram realizadas duas triagens.

A primeira triagem, foi aplicada para considerar apenas artigos de congresso, dissertações e teses. Após essa etapa, o número de registros encontrados foi de 523.

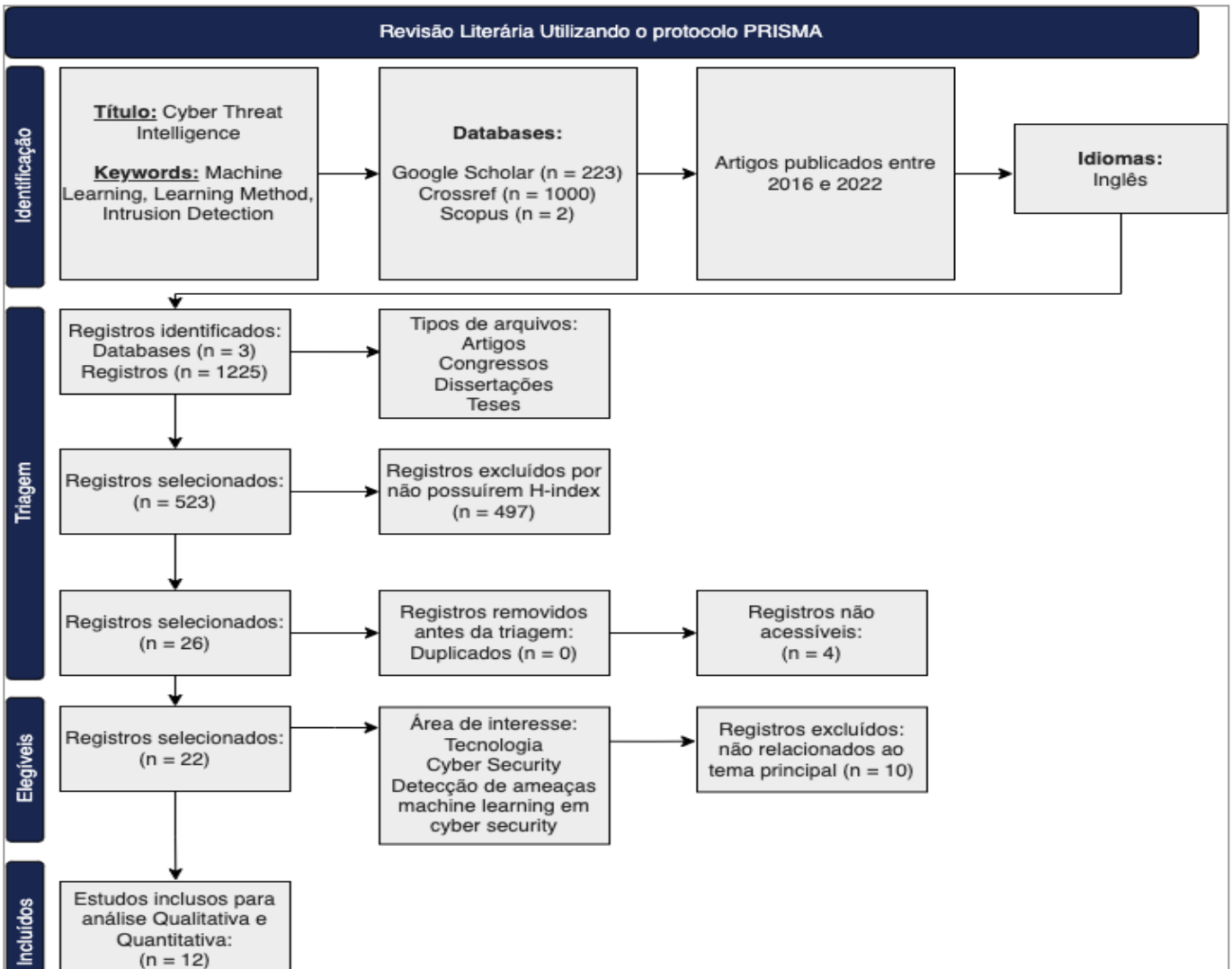
Na triagem seguinte, foi utilizado uma medição baseada em H-Index (*H-Index, 2021*), considerando apenas obras que são consideradas relevantes, possuindo um número proporcional de citações a assim categorizado como um *H-Index*. Com esta segunda triagem, foram removidos 497 registros. Trazendo um total de 26 documentos elegíveis para análise qualitativa.

Ao tentar acessar os documentos elegíveis, foram encontrados 4 artigos inacessíveis. Dos 22 documentos restantes, 10 não se enquadram na proposta desta pesquisa, não sendo relacionados a *Cybersecurity, Tecnologia, Detecção de ameaças ou Machine learning aplicado a Cybersecurity* . Totalizando assim 12 artigos disponíveis para análise qualitativa e quantitativa.

O resultado apresentado pode ser ilustrado baseando-se no protocolo PRISMA, conforme observado na

Figura 6.

Figura 6 - Protocolo PRISMA



Fonte: Resultado da Pesquisa

Os artigos selecionados para análise qualitativa e quantitativa podem ser observados na Tabela 1.

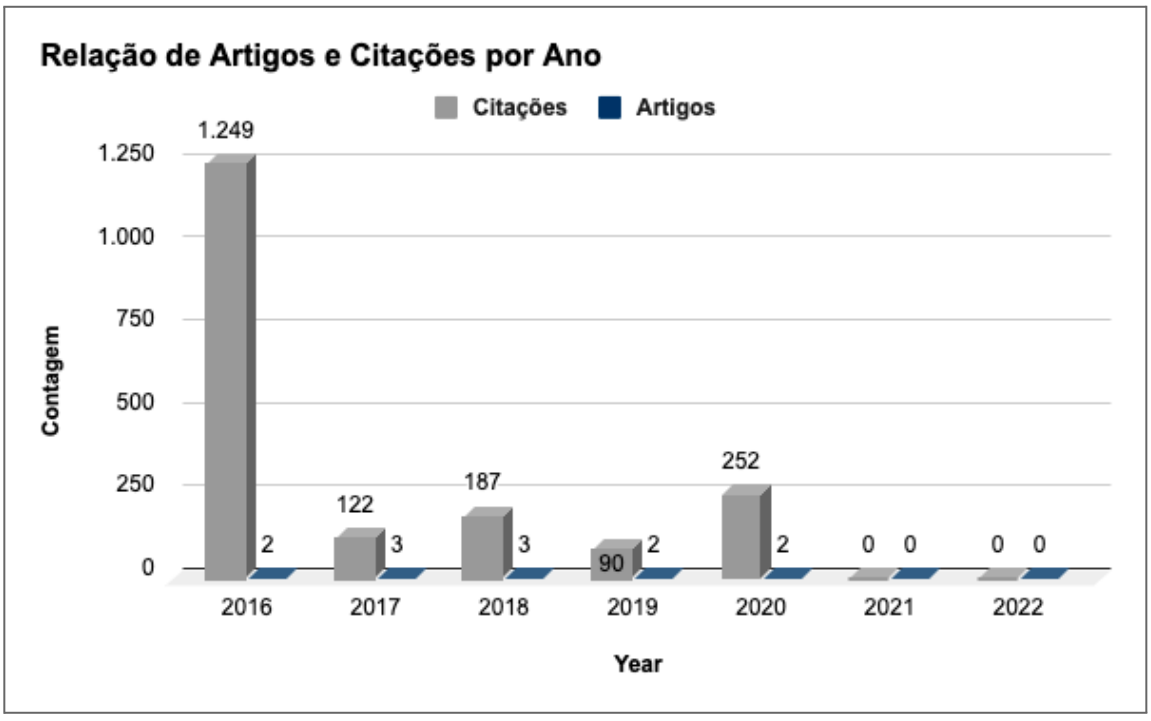
Tabela 1 - Relação de Artigos Selecionados

Authors	Title	Year
Anna L. Buczak, Erhan Guven	A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection	2016
Khurum Nazir Junejo, Jonathan Goh	Behaviour-Based Attack Detection and Classification in Cyber Physical Systems Using Machine Learning	2016
Charles Feng, Shuning Wu, Ningwei Liu	A user-centric machine learning framework for cyber security operations center	2017
Zecheng He, Tianwei Zhang, Ruby B. Lee	Machine Learning Based DDoS Attack Detection from Source Side in Cloud	2017
Tara Salman, Deval Bhamare, Aiman Erbad, Raj Jain,	Machine Learning for Anomaly Detection and Categorization in Multi-Cloud Environments	2017
George Loukas, Tuan Vuong, Ryan Heartfield, Georgi	Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning	2018
Mohit Sewak, Sanjay K. Sahay, Hemant Rathore	Comparison of Deep Learning and the Classical Machine Learning Algorithm for the Malware Detection	2018
Gozde Karatas, Onder Demir, Ozgur Koray Sahingoz	Deep Learning in Intrusion Detection Systems	2018
Umara Noor, Zahid Anwar, Tehmina Amjad, Kim-Kwa	A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise	2019
Ünal Çavuşoğlu	A new hybrid approach for intrusion detection using machine learning methods	2019
Mohamed Amine Ferrag, Leandros Maglaras, Sotiris	Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study	2020
Iqbal H. Sarker, Yoosef B. Abushark, Fawaz Alsolami,	IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model	2020

Fonte: Resultado da Pesquisa

A partir da análise bibliográfica, foi possível filtrar as publicações que podem ser consideradas relevantes, devido ao número de citações no período de 2016 a 2022. Foram publicados 2 artigos e 1249 citações em 2016, 3 artigos e 122 citações em 2017, 3 artigos e 187 citações em 2018, 2 artigos e 90 citações em 2019, 2 artigos e 252 citações em 2020. Não foram encontradas publicações relevantes no período de 2021 e 2022, até a data da pesquisa. A distribuição de artigos e citações por ano, pode ser observado no Gráfico 1.

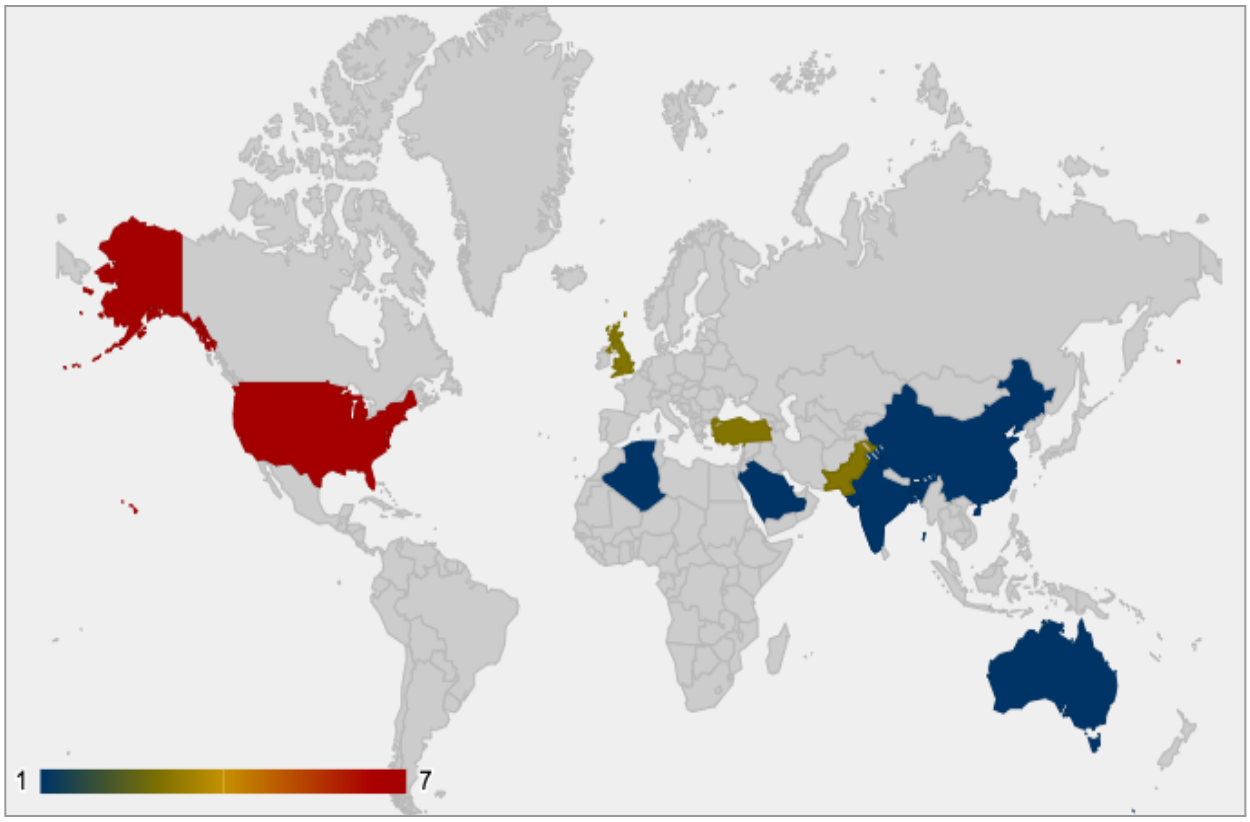
Gráfico 1 - Relação de Artigos Relevantes e Citações no período de 2016 a 2022.



Fonte: Resultado da Pesquisa

Na Figura 7 é possível observar a distribuição das publicações relevantes por país. Para isso foram consideradas as origens de instituições participantes em cada um dos artigos relevantes.

Figura 7 - Distribuição de Artigos Relevantes por País



Fonte: Resultado da Pesquisa

Analisando a figura, é possível observar a presença de publicações relevantes em quase todos os continentes, com exceção da América do Sul. Os Estados Unidos lidera em número de instituições envolvidas na construção de publicações relevantes, com 7 instituições participantes. Seguindo então, encontram-se Paquistão, Reino Unido e Turquia, todos com 3 instituições participantes, e por fim observa-se a participação de 1 instituição em cada um dos países: Algeria, Arábia Saudita, Austrália, Bangladesh, China, Cingapura, Índia e Qatar.

Utilizando como base os 12 artigos relevantes obtidos, foi realizada uma análise qualitativa e quantitativa em cada um dos mesmos. A análise levou em consideração, principalmente, o levantamento de técnicas de machine learning aplicadas na detecção e/ou classificação de agentes maliciosos, proposta de novos algoritmos e frameworks que se utilizem de aprendizagem de máquina para facilitar processos no domínio de Cyber Security.

Na sequência, apresentam-se o conteúdo analítico dos 12 artigos:

1. *A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, (BUCZAK;GUVEN,2016)* apresenta uma análise bibliográfica sobre a aplicabilidade de técnicas de machine learning e data mining para detecção de invasões em Cyber Security, assim como importantes datasets para treinamento de tais modelos. Em conclusão os autores apresentam a dificuldade de enquadrar as duas áreas, levantando o fato de que a apesar de efetivos, modelos de machine learning e data mining ainda não foram bem estabelecidos no campo de Cyber Security, devido a grande complexidade e métricas que devem ser analisadas para cada caso. Outro ponto importante levantado pelos autores é que além das técnicas, a fonte primordial para construção de qualquer modelo, são os dados, os quais foram identificados gaps na literatura de datasets voltados para cybersecurity, principalmente na falta de dados categorizados para treino de tais modelos. Como campo de pesquisas futuras, é apresentado a investigação de algoritmos de aprendizado incremental rápido, que possam beneficiar e impulsionar modelos de machine learning mal regulados e também auxiliar na detecção de ameaças.
2. *Behaviour-Based Attack Detection and Classification in Cyber Physical Systems Using Machine Learning,* (JUNEJO;GOH,2016) propõe a utilização de modelos de machine learning em Cyber-physical systems (CPS) para identificação de invasões, baseadas no comportamento das mesmas. Para isso, os autores exploram o treino de modelos de machine learning na detecção de ataques em maquetes de SWaTs (Secure Water Treatment). Dentro os modelos, observou-se que três algoritmos se destacam no experimento proposto. Os algoritmos de Random Forest (RF) e K-Nearest Neighbors (KNN, também conhecido como IBK no software Weka), apresentam uma alta taxa de recall, priorizando assim a baixa taxa de falsos negativos em comparação a falsos positivos. Já o algoritmo Best-First Tree (BFTree) apresentou a melhor taxa de acurácia e precisão. Em conclusão o BFTree pode ser escolhido em casos nos quais a acurácia deve ser priorizada, KNN quando o número de detecções for uma prioridade, mesmo que tempo de conclusão seja maior, e o RF quando a aplicado a cenários onde a tolerância ao falso alarme é mínima.

3. *A User-Centric Machine Learning Framework for Cyber Security Operations Center,* (FENG; WU;LIU,2017) apresenta um sistema de detecção de ameaças baseado em machine learning em detrimento a sistemas baseados em regras utilizados em SOCs (Security Operation Centers). Os autores descrevem o problema da alta taxa de falsos positivos em SOCs e o quanto isso pode sobrecarregar uma operação de monitoramento desses ambientes. Para o experimento, foram utilizados alguns algoritmos de machine learning, sendo estes: Multi-Layer Neural Networks (MNN), Random Forest (RF), Support Vector Machine (SVM) e Logistic Regression (LR), utilizando-se do ecossistema de big data que é gerado em operações de um SOC, como logs de segurança, informações de alerta e insights de analistas. Na análise foi possível observar que mesmo utilizando-se labels desbalanceados e limitados, algoritmos simples de machine learning, conseguem performar até cinco vezes melhor do que um sistema tradicional baseado em regras. Observou-se que entre os algoritmos de machine learning citados, MNN e RF demonstraram melhor desempenho.
4. *Machine Learning Based DDoS Attack Detection from Source Side in Cloud*, (HE;LEE,2017) o artigo tem como proposta a criação de um sistema para detecção de ataques DDOS para prevenir ataques em repositórios alocados em ambientes na nuvem. Durante o desenvolvimento da proposta, é citado que as medidas tomadas em relação a ataques DDOS, geralmente são passivas e efetuadas após o ataque já ter tomado efeito no sistema, tendo assim um tempo de resposta lento a ameaças e de difícil rastreabilidade. Para a construção do sistema de detecção, foram testados 9 algoritmos de machine learning, sendo Support Vector Machine (SVM) Linear Kernel o que demonstrou melhor acurácia e F1-Score, respectivamente 99.75% e 99.73%. Outros quatros algoritmos (SVM com Linear Poly Kernel, Decision Trees e Random Forest) demonstraram acurácia acima de 99%. Como desenvolvimento futuro do trabalho, é indicado pelos autores a combinação de diferentes algoritmos de machine learning, especialmente explorando algoritmos de aprendizagem não supervisionada.
5. *Machine Learning for Anomaly Detection and Categorization in Multi-Cloud Environments,* (SALMAN et al.,2017) apresenta a usabilidade de algoritmos de machine learning para detecção de anomalias e categorização de ataques em ambientes Multi-Cloud. A pesquisa demonstrou que algoritmos de aprendizado supervisionado como Linear Regression (LR) e Random Forest (RF), possuem uma alta acurácia. O algoritmo de Random Forest, em conjunto com processos de feature engineering, apresentou uma acurácia de detecção de anomalias próxima de 99%, e uma acurácia próxima de 93.6% para classificação dos ataques. Em conclusão, os autores citam que os algoritmos citados podem ser utilizados em ambientes Multi-Cloud para detecção e classificação de ataques, com mudanças sutis em sua modelagem, dado que apesar da alta taxa de acerto, os algoritmos ainda apresentam dificuldades para diferenciar certos tipos de ataque, devido a similaridade entre os comportamentos de tráfego, e que mais testes podem ser desenvolvidos utilizando mais dados e técnicas para diferenciação de features entre os tipos de ataques.
6. *Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning,* (LOUKAS et al.,2018) propõe a utilização de um sistema de detecção de invasões para veículos, baseado em modelos de Deep Learning. Os autores mostram neste estudo que a utilização de Recurrent Neural Networks (RNN) em conjunto com Long short-term memory (LSTM) apresentam um alto nível de acurácia para detecção de ameaças em veículos robóticos, superando em comparação, outros algoritmos clássicos de machine learning. O fato de que a maior desvantagem dos algoritmos de deep learning para aplicação em veículos robóticos seja a latência de detecção devido a demanda de processamento, pode ser mitigada utilizando-se da computação em nuvem e sua auto-escalabilidade. Os autores também citam ressalvas no experimento, ressaltando que existem diferenças na aplicação do experimento (realizado em um veículo robótico para teste) para casos reais e não controlados, devido à conexão com a internet, o que abre uma brecha para latência de conexão e outras possíveis ameaças.

7. *Comparison of Deep Learning and the Classical Machine Learning Algorithm for the Malware Detection,* (SEWAK;SAHAY;RATHORE,2018) realiza uma análise comparativa entre modelos de machine learning e deep learning para detecção de Malwares, com ênfase nos algoritmos de Random Forest (RF) e Deep Neural Networks (DNNs). Em conclusão a análise aponta que a utilização do Deep Neural Networks (DNNs) para detecção de Malwares não traz um bom custo benefício, dado que a abordagem tradicional de machine learning (Random Forest) trouxe um melhor resultado (99.78% em comparação a 99.21% do DNNs). Os autores ressaltam a necessidade de maiores análises para outras técnicas de deep learning, tais como Recurrent Neural Networks (RNN), Long short-term memory (LSTM) e Echo State Networks (ESN).
8. *Deep Learning in Intrusion Detection Systems,* (KARATAS;DEMIR,2018) apresenta em formas gerais as diferenças entre machine learning e deep learning, assim como o panorama de quais métodos e datasets são utilizados com maior frequência no construção de sistemas detectores de invasões. No trabalho comparativo é citado o Support Vector Machine (SVM) como um dos métodos mais disseminados no domínio de machine learning, e que, apesar de algoritmos de deep learning apresentarem grande eficiência para resolução de problemas, apresentam contrastes negativos em relação aos métodos tradicionais, dado o alto nível de consumo de processamento, quantidade de dados e tempo necessário para treino dos modelos. Como proposta de desenvolvimento futuro, os autores recomendam o desenvolvimento de testes para algoritmos de aprendizagem, utilizando datasets que contenham dados de invasões mais recentes.
9. *A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise,* (NOOR et al.,2019) propõe um framework para atribuição de CTAs (Cyber Threat Actors), utilizando modelos de machine learning treinados a partir de IOCs (Indicators of Compromise) de alto nível, em conjunto com dados do ATT&CK e MITRE, para classificação de ameaças cibernéticas em ambientes de FinTech. Em comparação aos métodos tradicionais de análise de regras usando somente relatórios de CTA, os modelos de machine learning performaram melhor em todas as métricas (precisão, recall, F-Score, false positive rate). Dentre os modelos apresentados (Naive Bayes, K-Nearest Neighbors, Decision Tree, Random Forest, e DLNN), o Deep Learning Neural Networks (DLNN), apresentou um desempenho geral melhor do que os outros algoritmos de machine learning.
10. *A new hybrid approach for intrusion detection using machine learning methods ,* (ÜNAL ÇAVUŞOĞLU,2019) apresenta uma proposta de algorítimo hibrido para detecção de invasões em detrimento a modelos tradicionais encontrados na literatura. O modelo proposto utiliza-se de métodos mais extensivos na seleção de features para detecção, para isso são usados mais outros modelos em conjunto para uma melhor seleção destes atributos, como Stacking Ensemble Learning, K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF) e Naive Bayes. Como resultado da comparação da nova proposta de modelo, o mesmo demonstrou ser mais performático na detecção de ataques do que muitos modelos encontrados na literatura, em todos os tipos de ataques.
11. *IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model ,* (SARKER et al., 2020) propõe um novo modelo de machine learning baseado em algoritmos de árvore, o Intrusion Detection Tree (IntruDTree) para detecção de invasões. Durante o desenvolvimento, o novo modelo é comparado com outros modelos tradicionais de machine learning utilizados para detecção de de invasões como Linear Regression (LR), K-Nearest Neighbors (KNN) e Support Vector Machines (SVM). O IntruDTree, é construído a partir da seleção e ranqueamento de features importante em Cyber Security e então implementado como base em um algoritmo tradicional de árvores. O novo modelo demonstrou maior acurácia, Recall e F-Score comparado aos modelos tradicionais, garantindo assim uma maior efetividade para detecção de novos casos e eficiência na economia de processamento, demandando um menor número de features para treinamento do modelo. Como trabalho posterior, os autores discutem a possibilidade de testar o IntruDTree com

datasets de maior escala, coletados a partir de IoT e assim verificar a sua efetividade e aplicabilidade no domínio de cybersecurity.

12. *Deep Learning for Cyber Security Intrusion Detection: Approaches,Datasets, and Comparative Study,* (FERRAG et al.,2020) os autores focam na análise de sistemas de detecção de invasões baseados em Deep Learning, assim como a categorização e avaliação de datasets comumente utilizados para construção de tais modelos. A divisão da análise se dá em sete categorias de modelos de deep learning, sendo elas: Neural Networks, Deep Neural Networks, Restricted Boltzmann Machine, Deep Belief Networks, Convolutional Neural Networks, Deep Boltzmann Machines e Deep Autoencoders. Descrevendo aplicações para cada uma das abordagens, nota-se a eficiência e alta acurácia desse tipo de modelo, cada qual atendendo a um tipo de cenário específico. Das análises realizadas, pode-se observar uma grande eficiência dos modelos de Convolutional Neural Networks para classificação de ataques, tendo como base o treino deste tipo de modelo com o dataset Contagio-CTU-UNB obteve-se uma acurácia de classificação em torno de 99%.

4. ANÁLISE DE RESULTADOS

Com base nos artigos relevantes, é possível realizar uma tabulação dos modelos utilizados nos trabalhos propostos, como observado na Tabela 2. Na tabulação são considerados todos os modelos de machine learning que foram aplicados em cada um dos artigos, incluindo também propostas de novos algoritmos.

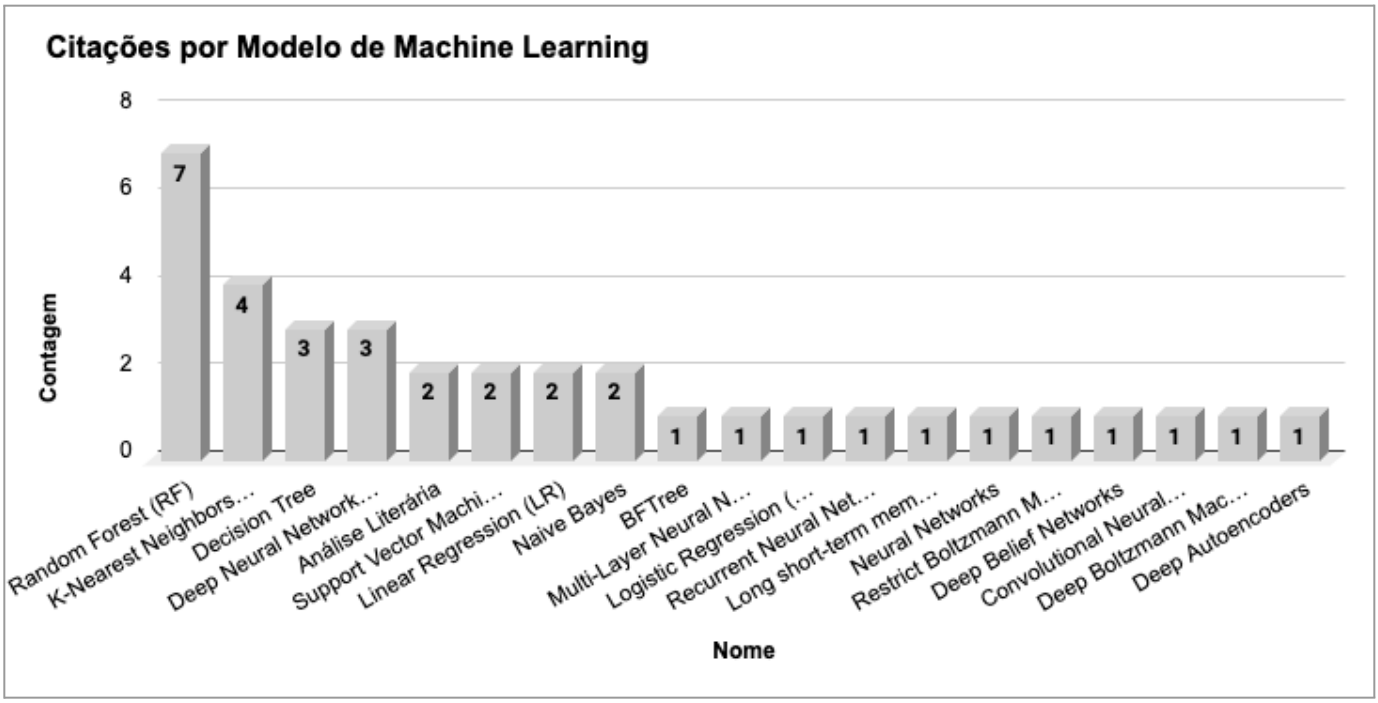
Tabela 2 - Modelos de Machine Learning utilizados para detecção de ameaças

Artigo	Ano	Modelos Apresentados						
1	2016	Análise Literária						
2	2016	BFTree	Random Forest (RF)	K-Nearest Neighbors (KNN)				
3	2017	Support Vector Machine (SVM)	Random Forest (RF)	Multi-Layer Neural Networks	Logistic Regression (LR)			
4	2017	Support Vector Machine (SVM)	Random Forest (RF)	Decision Tree				
5	2017	Linear Regression (LR)	Random Forest (RF)					
6	2018	Recurrent Neural Networks (RNN)	Long short-term memory (LSTM)					
7	2018	Deep Neural Networks (DNN)	Random Forest (RF)					
8	2018	Análise Literária	-					
9	2019	Naive Bayes	Random Forest (RF)	K-Nearest Neighbors (KNN)	Decision Tree	Deep Neural Networks (DNN)		
10	2019	Naive Bayes	Random Forest (RF)	K-Nearest Neighbors (KNN)	Decision Tree			
11	2020	Support Vector Machines (SVM)	K-Nearest Neighbors (KNN)	Linear Regression (LR)				
12	2020	Neural Networks	Restrict Boltzmann Machine	Deep Belief Networks	Convolutional Neural Networks	Deep Boltzmann Machines	Deep Autoencoders	Deep Neural Networks (DNN)

Fonte: Resultado da Pesquisa

Baseando-se na tabulação dos modelos de machine learning citados nos artigos, é possível criar um histograma, evidenciando o número de vezes que cada uma das abordagens foi utilizada, conforme observa-se no Gráfico 2.

Gráfico 2 - Quantidade de Citações por modelo de Machine Learning

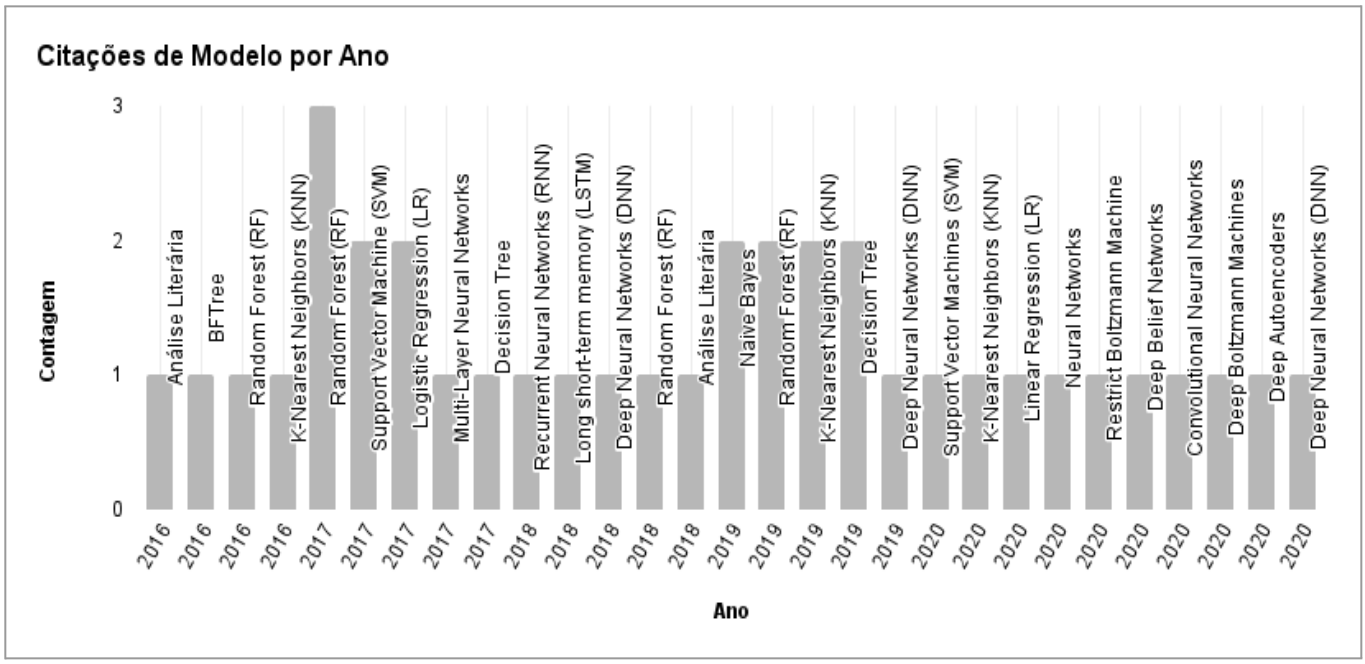


Fonte: Resultado da Pesquisa

Analisando o Gráfico 2, é possível verificar a que Random Forest (RF) é o modelo mais citado dentre os artigos relevantes, possuindo quase o dobro de frequência do que o segundo modelo mais utilizado, sendo este o K-Nearest Neighbors. O Decision Tree aparece como terceiro algoritmo mais utilizado junto ao Deep Neural Networks (DNNs).

A partir da tabulação também é possível verificar quais são os modelos mais citados nos artigos relevantes por ano, como pode ser observado no Gráfico 3.

Gráfico 3 - Citações de Modelo por Ano



Fonte: Resultado da Pesquisa

Analisando o Gráfico 3, observa-se que o Random Forest apresenta uma boa distribuição, sendo citado em todos os anos, com exceção de 2020, sendo em 2017 o algoritmo mais citado e nos outros anos também permanece com um dos mais citados. Após 2016 em todos os anos foi citado algum algoritmo de Deep Neural Networks (DNNs), não sendo o mais citado em números gerais, porém mantendo uma frequência importante em comparação aos demais algoritmos. O K-Nearest Neighbors e Decision Trees não aparecem de forma constante em citações ao longo dos anos, apesar de apresentarem um número relevante de citações totais, diferente do Random Forest e Deep Neural Networks que apresentam uma frequência constante.

O Random Forest possui uma alta relevância devido a sua versatilidade, podendo ser utilizado tanto para regressões quanto para classificações. Como observado em ambientes multi-cloud, o algoritmo demonstrou uma acurácia de 99% para detecção de anomalias (SALMAN et al.,2017), apresentando um desempenho superior a algoritmos mais robustos na detecção de Malwares, tendo 99.78% de acurácia em comparação a 99.21% do Deep Neural Networks (DNNs) (SEWAK;SAHAY;RATHORE,2018).

O funcionamento do algoritmo também contribui para que ele trabalhe bem com datasets que incluam noisy data, o que favorece a aplicação em datasets clássicos utilizados para treinamento de modelos de machine learning referente a ataques e atores de ameaça como KDD e também datasets mais novos e que incluem novos tipos de dados para categorização como UNSW (SALMAN et al.,2017). A modelagem deste algoritmo, garante que no processo de categorização, ele possua um alto desempenho em cenários onde possam possam haver sobreposição de características, principalmente na identificação de ameaças cibernéticas (NOOR et al,2019).
Outro fator importante para utilização do Random Forest, é a sua alta taxa de recall, priorizando assim a baixa taxa de falsos negativos em comparação a falsos positivos (JUNEJO;GOH,2016) o que pode ser importante em cenários de detecção de ameaça, onde o impacto de falsos negativos é muito maior do que falsos positivos.

O K-Nearest Neighbors (KNN) também possui uma alta versatilidade, também podendo ser utilizado para problemas de regressão, apesar de comumente ser utilizado para processos de classificação. Junto ao Random Forest, apresenta taxas relativamente inferiores para classificação e também prioriza uma baixa taxa de falsos negativos em comparação a falsos positivos. Devido ao approach de Lazy-Algorithm, o KNN possui um tempo maior para classificações em comparação a outros algoritmos, tendo problemas para escalar sua performance e demandando maior uso de memória. O seu uso com dados de alta dimensão, também tende a causar overfitting (IBM, 2022).

O KNN também possui variações, como Indexed Partial Distance Search kNearest Neighbor (IKPDS), algoritmo que utiliza-se da base do K-Nearest Neighbors em conjunto com outras técnicas de modelagem para trazer um mesmo nível de acurácia, porém reduzindo o tempo de processamento o qual o KNN é mais deficiente. Em aplicações voltadas para detecção de invasões em redes, o IKPDS manteve uma acurácia próxima a 99.6% e com um tempo de processamento para classificação relativamente menor (RAO;SWATHI, 2017).

O K-Nearest Neighbors é recomendado quando o tempo de processamento não for uma prioridade, mas sim o número de detecções, onde o algoritmo apresenta um desempenho relativamente maior (JUNEJO;GOH,2016).

Algoritmos de Deep Learning, como o Deep Neural Networks, que são um subconjunto de algoritmos clássicos de machine learning, utilizando-se de técnicas mais robustas para abstrair padrões implícitos nos dados e estabelecer regras de decisões mais eficientes, não performaram com um bom custo benefícios em suas aplicações. Apesar de ter apresentarem um bom desempenho para detecção de invasões, junto a utilização de outras técnicas como o Recurrent Neural Networks (RNN) e Long short-term memory (LSTM) (LOUKAS et al.,2018), este tipo de algoritmo acaba requerendo um maior volume de dados e consequentemente um maior poder computacional, o que acarreta não somente em um maior custo, como também maior dificuldade de treino, seleção de features e implementação (KARATAS;DEMIR,2018).

O Deep Neural Networks, apesar de apresentar um desempenho sutilmente inferior em acurácia para detecção de malwares em comparação ao Random Forest (SEWAK;SAHAY;RATHORE,2018), em outros cenários utilizando-se do Convolutional Neural Networks (CNNs), esse tipo de algoritmo, comumente utilizado para técnicas de visão computacional e possuindo ampla aplicação em setores de saúde, marketing e indústria automotiva (IBM, 2020) apresentou uma maior acurácia de classificação de malwares, utilizando o dataset Contagio-CTU-UNB, aproximando-se dos 99% de acurácia (FERRAG et al.,2020).

A junção de técnicas de machine learning também mostra-se eficiente, dado que a maior parte dos problemas no domínio de Cyber Security e posteriormente em processos de Cyber Threat Intelligence, baseiam-se na detecção e categorização do agente de ameaça para uma ação reativa e/ou proativa, sendo assim, é possível usar algoritmos que possuam melhor desempenho para cada um de seus domínios (classificação e regressão) e juntar os resultados para um melhor conjunto de desempenho, como observado em aplicações de detecção de invasões, onde é necessário detectar e classificar ameaças de forma proativa (ÜNAL ÇAVUŞOĞLU,2019).
A construção de novos algoritmos baseados na junção de algoritmos clássicos de machine learning, utilizando-se dos pontos de maior eficiência de cada um deles, também pode trazer melhorias em Recall, F-Score e aumentar a efetividade de detecção de novos casos de invasão e reduzir a complexidade de processamento, demandando um processo mais sucinto de feature engineering e implementação (SARKER et al., 2020).

Com base nos artigos analisados também é possível observar que uma variedade de nichos podem se beneficiar da utilização de machine learning para o domínio de Cyber Security. Entre os artigos levantados, foram explorados cenários levantando em conta setores como o de FinTechs, carros inteligentes, instalações físicas e ambientes multi-cloud.

5. CONSIDERAÇÕES FINAIS

Este trabalho realizou uma análise textual, utilizando como base um livro considerado como referência para a utilização de métodos de machine learning para Cyber Security. Posteriormente foi realizada uma análise bibliométrica, utilizando-se de três databases disponíveis e utilizando palavras chaves encontradas. Com base nos artigos relevantes encontrados, foi realizado uma análise qualitativa dos artigos, para encontrar como técnicas de machine learning são aplicadas no domínio de Cyber Security, quais são as possibilidades de aplicação e seus benefícios para tomadas de decisão em diferentes abordagens e que possam beneficiar no processo de Cyber Threat Intelligence.

Como principal contribuição, este trabalho traz consigo a identificação de quais os principais modelos e técnicas de machine learning utilizados no domínio de Cyber Security, afirmando a proposição inicial de que o modelo de Random Forest é uma técnica predominante utilizada na literatura. Modelos mais complexos como Deep Learning, apesar de apresentar resultados relativamente maiores em acurácia, acabam não trazendo um bom custo benefício, dado a necessidade de um maior volume de dados e maior poder computacional para implementação.

Na análise qualitativa de cada um dos artigos relevantes levantados, foi possível verificar a aplicação de machine learning de diferentes maneiras e em assuntos diversos, como FinTechs, carros inteligentes, instalações físicas, ambientes multi-cloud, detecção de Malwares e detecção de ataques DDOS.

O estudo realizado apresenta limitações quanto ao número de bases utilizadas, restringindo a bibliometria em 3 bases: Crossref, Google Scholar e Scopus. O número de modelos observados também limita-se aos 12 artigos relevantes que foram considerados elegíveis para a análise.

Nota-se que o tema está em grande ascensão e possui diversas aplicações para ambientes corporativos e governamentais, dado que que modelos de machine learning podem auxiliar na tomada de decisão e

melhor entendimento das ameaças, tanto em sua detecção, quanto na sua categorização. Isto propicia um processo de Cyber Threat Intelligence, baseado em dados e potencializando a interação humano computador a fim de detectar agentes maliciosos de forma mais eficiente.

Como trabalhos futuros, pode-se realizar uma demonstração comparativa entre o Random Forest, K-Nearest Neighbors e Deep Neural Networks, aplicados a etapas de Cyber Threat Intelligence, a fim de avaliar o custo benefício, principais métricas, velocidade e adaptação a novos datasets que incluam dados mais recentes.

REFERÊNCIAS

CISA. Federal Virtual Training Environment. Disponível em: <https://fedvte.usalearning.gov/publiccourses/ici/iciframe.php>. Acesso: 01 nov. 2022.

USA JOINT INTELLIGENCE. International Registration Plan. Disponível em: <https://www.ncsc.gov.uk/files/An-introduction-to-threat-intelligence.pdf>. Acesso: 01 nov. 2022.

NATIONAL CYBER SECURITY CENTER. An introduction to threat intelligence. Disponível em: <https://www.ncsc.gov.uk/files/An-introduction-to-threat-intelligence.pdf>. Acesso: 01 nov. 2022.

CISA. Assessing the potential value of Cyber Threat Intelligence (CTI). Disponível em: <https://www.cisa.gov/sites/default/files/publications/Assessing%20Cyber%20Threat%20Intelligence%20Threat%20Feeds_508c.pdf>. Acesso: 01 nov. 2022.

R. M. Lee. Intelligence Defined and its Impact on Cyber Threat Intelligence. Disponível em: <http://www.robertmlee.org/tag/intelligence/>. Acesso: 01 nov. 2022.

MIT SLOAN SCHOOL OF MANAGEMENT. Machine Learning Explained. Disponível em: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. Acesso: 01 nov. 2022.

ANALYTICS VIDHYA. Understanding Random Forest. Disponível em: <https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>. Acesso: 01 nov. 2022.

CORPORATE FINANCE INSTITUTE. Random Forest. Disponível em: <https://corporatefinanceinstitute.com/resources/data-science/random-forest/#:~:text=Advantages%20of%20Random%20Forests&text=Missing%20values%20are%20substituted%20by,numerous%20variables

%20running%20into%20thousands.>. Acesso: 01 nov. 2022.

IBM. Random Forest. Disponível em: <https://www.ibm.com/cloud/learn/random-forest>. Acesso: 01 nov. 2022.

SENTINEL ONE. Advancing Security | The Age of AI & Machine Learning in Cybersecurity. Disponível em: <https://www.sentinelone.com/blog/advancing-security-the-age-of-ai-machine-learning-in-cybersecurity/#:~:text=Threat%20Operations%20Management%20%E2%80%93%20AI%20technology,and%20speed%20of%20the%20attacker>. Acesso: 01 nov. 2022.

OUTPOST24. What are AI and machine learning adding to threat intelligence – brains, brawn or both?. Disponível em: <https://outpost24.com/blog/what-are-ai-and-machine-learning-adding-to-threat-intelligence-brains-brawn-or-both>. Acesso: 01 nov. 2022.

SPLUNK. What is Cyber Threat Intelligence?. Disponível em: <https://www.splunk.com/en_us/data-insider/threat-intelligence.html>. Acesso: 01 nov. 2022.
CYBERCRIME MAGAZINE. Cybercrime To Cost The World $10.5 Trillion Annually By 2025. Disponível em: <https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/>. Acesso: 01 nov. 2022.

CISA. Cyber Security Awareness Month 2021. Disponível em: <https://www.cisa.gov/sites/default/files/publications/Cybersecurity%20Awareness%20Month%202021%20-%20Why%20is%20Cybersecurity%20Important.pdf>. Acesso: 01 nov. 2022.

IBM. What is the k-nearest neighbors algorithm?, 2020. Disponível em: <https://www.ibm.com/topics/knn>. Acesso: 01 dez. 2022.

IBM. Convolutional Neural Networks. Disponível em: <https://www.ibm.com/cloud/learn/convolutional-neural-networks>. Acesso: 01 dez. 2022.

IBM. Deep Learning, 2020. Disponível em: <https://www.ibm.com/cloud/learn/deep-learning>. Acesso: 04 dez . 2022.

IBM. What is a Decision Tree, 2020. Disponível em: <https://www.ibm.com/cloud/learn/deep-learning>. Acesso: 04 dez . 2022.

KDNuggets. Decision Tree Algorithm, Explained, 2022. Disponível em: <https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html>. Acesso: 04 dez. 2022.

University of Wisconsin. STAT 479: Machine Learning, 2018. Disponível em: <https://sebastianraschka.com/pdf/lecture-notes/stat479fs18/02_knn_notes.pdf>. Acesso: 04 dez. 2022.

SMEBOOK. The advantages and disadvantages of deep learning, 2021. Disponível em: <https://smebook.eu/knowledge-base/deep-learning/the-advantages-and-disadvantages-of-deep-learning>. Acesso: 04 dez. 2022.

DATTO. 5 Amazing Applications of Deep Learning Cybersecurity, 2022. Disponível em:

<https://www.datto.com/blog/5-amazing-applications-of-deep-learning-in-cybersecurity#:~:text=Deep%20learning%2C%20convolutional%20neural%20networks,bad%20and%20good%20network%20activities.>. Acesso: 04 dez. 2022.

AWS. O que é overfitting?, 2022. Disponível em: <https://aws.amazon.com/pt/what-is/overfitting/>. Acesso: 08 dez. 2022.

KDNuggets. More Performance Evaluation Metrics for Classification Problems You Should Know. 2022. Disponível em: <https://www.kdnuggets.com/2020/04/performance-evaluation-metrics-classification.html>. Acesso: 08 dez. 2022.

ALSAEDI et al. Cyber Threat Intelligence-Based Malicious URL Detection Model Using Ensemble Learning. MDPI, 2022.

BUCZAK; GUVEN. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE, 2016.

FENG;WU;LIU. A User-Centric Machine Learning Framework for Cyber Security Operations Center. IEEE,2017.

FERRAG et al. Deep Learning for Cyber Security Intrusion Detection: Approaches,Datasets, and Comparative Study. Science Direct, Journal of Information Security and Applications, 2020.

HE;LEE. Machine Learning Based DDoS Attack Detection from Source Side in Cloud. IEEE, 2017.

LIAO;VEMURI. Use of K-Nearest Neighbor classifier for intrusion detection. Elsevier, 2002.

JUNEJO;GOH. Behavior-Based Attack Detection and Classification in Cyber Physical Systems Using Machine Learning. Association for Computing Machinery, 2016.

KARATAS;DEMIR. Deep Learning in Intrusion Detection Systems. IEEE, 2018.

LOUKAS et al. Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning. IEEE, 2018.

MARKEY. Using Decision Tree Analysis for Intrusion Detection. SANS Institute, 2011.

NOOR et al. A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise. Elsevier, FGCS, 2019.

PLANQUE. Cyber Threat Intelligence From confusion to clarity; An investigation into Cyber Threat Intelligence. Leiden University Student Repository, 2017.

QIANG et al. Framework of Cyber Attack Attribution Based on Threat Intelligence. Springer, 2017.
SALMAN et al. Machine Learning for Anomaly Detection and Categorization in Multi-Cloud Environments. IEEE, 2017.

RAO;SWATHI. Fast kNN Classifiers for Network Intrusion Detection System. Indian Journal of Science

and Technology,2017.

SARKER et al. IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model. Symmetry, MDPI, 2020.

SEWAK;SAHAY;RATHORE. Comparison of Deep Learning and the Classical Machine Learning Algorithm for the Malware Detection. IEEE, 2018.

ÜNAL ÇAVUŞOĞLU. A new hybrid approach for intrusion detection using machine learning methods. Springer, 2019.