

Aplicações de machine learning em IoT: uma visão sobre o estado da arte

João Vitor Maia Neves Cordeiro

21 de março de 2021

Resumo

O presente trabalho visa fomentar um debate sobre o estado da arte das aplicações de machine learning (ML) em tecnologias Internet of Things (IoT), a fim de dar ao leitor o material necessário para compreender a importância da multidisciplinaridade entre as mais diversas áreas da computação; O artigo está dividido da seguinte forma: uma breve introdução sobre os temas ML e IoT; uma série de exemplos práticos da interligação entre as duas áreas estudadas; e por fim uma discussão sobre o que há de mais avançado sendo desenvolvido em universidade e indústrias pelo mundo; Como metodologia para a coleta de dados foi realizada uma pesquisa bibliográfica buscando materiais publicados nos últimos 3 anos em periódicos e congressos nacionais e internacionais.

1 Introdução

1.1 Motivação

O crescente interesse nos estudos da área de machine learning vem acompanhado de uma certa confusão quanto as reais aplicação dessa tecnologia, sendo disseminada principalmente por publicações equivocadas da grande mídia [1]. Desta forma, é de extrema importância que a comunidade científica formule publicações que tenham como objetivo colaborar com a formação de uma melhor fonte de conhecimento para o público geral.

Da mesma forma, é notável que a expansão do mercado de IoT seja suportada pelo avanço das tecnologias em redes de computadores e mais recentemente pelo aprendizado de máquina.

1.2 Justificativa

Como será visto melhor em seguida, os dois principais conceitos trabalhados nesse artigo estão em grande foco no cenário mundial, com novas pesquisas sendo realizadas todos os dias tanto no âmbito acadêmico quanto institucional.

Dessa forma, é essencial que sejam realizadas revisões bibliográficas sistemáticas com frequência, para verificar qual o estado atual dessas pesquisas. Somente a partir de um processo de revisão desses é possível compreender o estado da arte e definir prioridades de pesquisa futuras.

1.3 Objetivos Gerais

Esse artigo se propõe a introduzir alguns conceitos básicos para o melhor entendimento de seu conteúdo, discutir sobre as principais aplicações de técnicas de aprendizado de máquina em IoT bem como apontar avanços notáveis nessa área.

1.4 Objetivos Específicos

- Introduzir ao leitor os assuntos machine learning e IoT.
- Listar de forma não extensiva as aplicações possíveis de técnicas de machine learning em dispositivos IoT
- Comentar sobre o estado da arte das já citadas aplicações, com situações reais.

1.5 Organização do artigo

Este artigo está organizado da seguinte forma. Na seção 2, conceitos básicos, temos uma breve introdução aos conceitos que serão necessários para o entendimento do trabalho, a definição e explicação de IoT e Machine Learning; Na seção 3, trabalhos correlatos estão outras obras relacionadas da área, a fim de explicar a importância das aplicações de machine learning em dispositivos

IoT; Na seção 4, aspectos relevantes são permeados os assuntos importantes para a contextualização do texto. Além destes, a seção também trabalha pontos específicos que devem ser levados em consideração quando tratamos de aplicações de machine learning em dispositivos IoT; Na seção 5, problemas existentes, aqui entramos nas dificuldades e desafios associados à esta área de pesquisa; Na seção 6, soluções possíveis, busca-se apresentar soluções para os desafios propostos na quinta seção; Na seção 9, referências bibliográficas estão listados todos os trabalhos utilizados como base para a confecção deste artigo, dando o devido crédito aos autores originais.

2 Conceitos básicos

2.1 Machine Learning

O termo machine learning foi cunhado por Samuel Lee Arthur, ainda na década de 50, quando Arthur trabalhava como pesquisador para a IBM desenvolvendo um programa capaz de aprender a jogar xadrez. O algoritmo em questão utilizava dados de outras partidas para prever a probabilidade de vitória a cada ação, tomando o caminho com a maior probabilidade de vitória. O termo se refere a capacidade dos algoritmos de ML de evoluírem conforme a execução, baseando-se em dados para aprender a melhor forma de realizar uma função.

De forma mais prática, machine learning é área da ciência da computação que estuda algoritmos que podem aprimorar seu funcionamento por base de experiências prévias, sejam essas fornecidas por um banco de dados ou pelos dados que o algoritmo gera durante sua execução. É extensivamente utilizada em situações onde algoritmos convencionais não podem resolver o problema de forma satisfatória ou em tempo hábil.

2.2 Internet of Things

A primeira definição do termo *internet of things (IoT)* veio do trabalho de Ashton *et al.* e pode ser traduzido para "uma rede aberta e abrangente de objetos inteligentes que têm a capacidade de se autoorganizar, compartilhar informações, dados e recursos, reagir e agir diante de situações e mudanças no ambiente". Desde então, os "objetos inteligentes" foram renomeados para "coisas" (*things*) e são descritos como dispositivos de hardware e software

capaz de se conectar a uma rede com capacidade para performar uma funcionalidade específica.

Hoje, mais de 20 anos após a primeira definição, os dispositivos IoT invadiram as casas e indústrias do mundo, adquirindo um papel essencial na sociedade moderna e que segundo recente análise [2] está em crescente tanto em lucro quanto em produtividade. Essas tecnologias podem se espalhar por diversas áreas, sendo as mais notáveis: automação residencial, agricultura inteligente e automação industrial.

3 Trabalhos correlatos

A tabela a seguir contendo uma revisão bibliográfica sistêmica realizada em buscas no Scholar Google nos mostra que há uma quantidade massiva de trabalhos realizados na área de IoT e machine learning, o que também é observado quando utilizamos as duas keywords simultaneamente.

Os resultados apontam que existe um interesse global nos assuntos, com a maioria dos trabalhos sendo escritos em inglês mas espalhados em países pelo mundo todo.

Keywords	Resultados
IoT	1,120,000
Machine Learning	4,670,000
IoT Machine Learning	176,000
Machine learning applications on IoT	161,000
Machine learning on sensors	1,380,000

3.1 A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security [3]

O artigo de Mohammed Ali Al-Garadi *et al.* tem como premissa a quantidade crescente de dispositivos conectados que operam com pouca intervenção humana para questionar a importância de se preocupar em dobro com segurança em um contexto de IoT.

Os dispositivos IoT contam com necessidades específicas da área e muitas vezes únicas quanto ao escopo de desenvolvimento e aplicação, por consequência disso podem acabar deixando a desejar quando o assunto é segurança, sendo protegidas por meios comuns e menos modernos de criptografia e controle de acesso.

Considerando isto, as técnicas de machine learning que obtiveram grandes avanços no passado recente, se mostram ferramentas capazes de suprir as necessidades de segurança da IoT. O autor mostra, de forma direta e acessível, informações sobre as tecnologias em machine learning aplicadas a segurança de dispositivos de IoT, criando uma base sólida de conhecimento para auxiliar no direcionamento de futuras pesquisas do assunto.

3.2 Crop Management with the IoT: An Interdisciplinary Survey [4]

O trabalho publicado por Vitali *et al.* começa com uma conceituação histórica sobre a lenta evolução tecnológica da agricultura até a revolução verde, que acelerou vertiginosamente a inovação na área. Após isso, os autores citam a introdução dos dispositivos IoT na agricultura como uma segunda revolução verde, junto com outras novas tecnologias que aos poucos são incorporadas nas lavouras, como *machine learning* e *edge computing*.

São citadas como principais causas dessa proliferação de tecnologias no campo o barateamento de constante de *chips* capazes de executar as aplicações de ML, o desenvolvimento de novos sensores capazes de desempenhar suas funcionalidades com menos gasto energético e o avanço de técnicas como o *deep learning*.

Os pesquisadores da universidade bolognesa ainda concluem apontando que a ampliação do uso dessas tecnologias em fazendas inteligentes é capaz de otimizar os lucros de proprietários de terras, assim como de diminuir impactos ambientais causados pelos plantios e aumentar a produção global de alimentos e insumos. Por fim, também fazem uma projeção de que em poucos anos essas tecnologias estarão nas mãos até de pequenos fazendeiros individuais continuando a melhorar a vida no campo.

3.3 Realizing an Effective COVID-19 Diagnosis System Based on Machine Learning and IOT in Smart Hospital Environment [5]

Uma das áreas ainda não citada aqui mas que possui grandes incentivos para pesquisas em IoT são os Smart Hospitals, ambientes de amparo a saúde controlados por dispositivos interconectados que podem performar diversas funções, desde a captação de sintomas até a administração de medicações.

Durante o ano de 2020, essas pesquisas tenderam-se a focar na pandemia do vírus SARS-COV-2, como é o caso do artigo citado.

No trabalho publicado por Abdulkareem *et al.* nós temos uma pesquisa de cunho mais experimental, lidando diretamente com amostras do vírus. Segundo o texto, foram coletados dados de pacientes ao redor do mundo com a ajuda de sensores em dispositivos inteligentes. Após isso, foram treinados modelos de 3 tipos: Naive Bayes (NB), Random Forest (RF), e Support Vector Machine (SVM). Dos 3 citados, o que se saiu melhor nos testes foi o modelo SVM, com uma precisão de até 95% em diagnósticos, sendo o sugerido pelos pesquisadores para uma possível aplicação em campo.

O artigo conclui que o modelo de ação proposto, que utiliza o Machine Learning como suporte para decisões clínicas ainda é superior aos modelos completamente automatizados, já que o fator humano continua sendo de grande importância para mitigar possíveis erros da máquina.

3.4 An architecture for adaptive task planning in support of IoT-based machine learning applications for disaster scenarios [6]

Todos os anos, milhares de desastres naturais e acidentes causados por humanos acontecem pelo mundo, pondo em risco a vida não só das pessoas afetadas mas também dos profissionais responsáveis por socorrer e resgatar as vítimas dessas situações. O uso de drones para cenários como esses se tornou popular recentemente, devido ao barateamento da tecnologia utilizada para o funcionamento básico dessas máquinas. O artigo publicado por Sacco *et al.* descreve a arquitetura de uma possível aplicação de machine learning no aprimoramento desses dispositivos.

Ao alimentar algoritmos de aprendizado de máquina com dados em vídeo captados por drones em desastres é possível treinar um modelo capaz de interpretar em tempo real a situação de um novo desastre e escolher a melhor ação a ser tomada nesse caso. Em específico, a arquitetura proposta pelo trabalho lida com um cenário de múltiplos agentes semi autônomos que são programados para realizar tarefas gerais como apagar chamas, remover escombros e carregar vítimas. Após treinado, o modelo é conectado ao drone e passa a monitorar as atividades que ocorrem no ambiente ao redor dele, pronto para identificar e responder a situações de risco.

A conclusão dos pesquisadores é de que apesar de promissor, o modelo

ainda conta com algumas barreiras, como a baixa precisão de sensores em situações tão extremas. É possível que com o avanço na tecnologia dos sensores, aplicações como essa possam se tornar 100% autônomas, salvando a vida de milhares de pessoas em desastres, sem arriscar a vida de socorristas.

4 Aspectos Relevantes

Durante a leitura dos artigos de base foi constatada uma interdisciplinaridade intensa, tanto em problemas quanto em soluções propostas a estes. É notável o esforço de pesquisadores da área da computação para compreender assuntos que tocam outras áreas do conhecimento e assim propor abordagens inteligentes para os dilemas que surgem.

Por se tratar de uma área em ascensão, o espaço disponível para novas ideias e novas tentativas se mostra enorme, considerando que para cada uma das áreas aplicadas é possível que um algoritmo X ou Y se dê melhor para a solução proposta, com novos algoritmos sendo criados todos os dias e evoluções sendo feitas nos algoritmos existentes [5].

Outra coisa que chamou atenção foi a maneira como a evolução constante e barateamento de sensores impacta nos futuros trabalhos a serem desenvolvidos na área, se hoje um equipamento se mostra inviável pelo custo de produção ou pela falta de precisão no sensor é possível que em anos ou mesmo meses, com essa barreira tecnológica sendo quebrada, algoritmos e aplicações de machine learning que se mostravam ineficientes acabem se tornando o estado da arte [6] [3].

Ao fim, outro aspecto importante a se citar é que pelos assuntos a serem tratados nesse artigo terem um tempo de vida relativamente curto, ainda estamos possivelmente no começo de uma revolução completa no modo de vida dos seres humanos, e apesar de aspectos éticos e de segurança ainda serem preocupações importantes, as tecnologias estudadas vem se mostrando com usos impressionantes para o auxílio no crescimento do nosso padrão de vida.

5 Problemas Existentes

Foi possível identificar durante a leitura dos trabalhos correlatos que existem problemas de certa forma estão interconectados, na maior parte limitações

de hardware. Nessa seção, esses problemas serão esmiuçados para futura análise e proposição de soluções. O primeiro problema a ser citado foi termo comum entre a maior parte das publicações analisadas, a relação entre custo dos sensores e a precisão das medições.

Algoritmos de machine learning precisam de dados extremamente precisos, tanto durante o treinamento do modelo quanto durante sua execução, tendo isso em vista é necessário que os sensores que irão captar os dados utilizados pelos modelos tenham uma precisão satisfatória. Entretanto, quanto mais preciso o sensor, mais pesado para o bolso de quem irá custear o projeto e a produção. É normal que no impulso para economizar nos sensores os fabricantes maiores de dispositivos IoT acabem negligenciando essa precisão, o que não pode de maneira nenhuma ocorrer em cenários de serviços críticos com aplicações de machine learning [6].

Outro problema relacionado ao hardware dos sistemas que é apontado por diversos autores é falta de poder computacional que a maior parte dos dispositivos IoT, o que faz com que os modelos tenham que ser rodados em servidores remotos com o auxílio de *cloud computing*, impactando no tempo de resposta (pois mesmo que a latência seja baixa, o tempo de upload e download dos dados já pode alterar significativamente a performance do algoritmo e do sistema como um todo). Por outro lado, dispositivos que possuem maior poder computacional, além de terem acréscimo no preço também precisam de mais energia para funcionar, aumentando o tamanho e custo dos dispositivos significativamente.

Outro problema conhecido e explorado por diversos autores é a preocupação quanto a segurança dos dispositivos, já que os mesmos possuem capacidades limitadas de criptografia e estão sujeitos a ataques maliciosos por utilizarem autenticações mais simples. O problema explorado por Ali *et al* [3] não possui ainda uma solução definitiva, tendo em vista de que ele surge da própria natureza e definição de um dispositivo IoT. Sendo assim, somente soluções parciais ou restritas a poucos casos de uso foram implementadas e pesquisadas até o momento.

6 Soluções Possíveis

Por tratar-se de uma área ampla, com diversos problemas que não foram citados na seção anterior por terem menor relevância ou possuírem pouco material base de pesquisa, aqui iremos apenas propor algumas soluções possíveis

para uma listagem não extensiva dos principais problemas do campo de estudo. Começaremos abordando o problema do poder de processamento citado acima, onde os dispositivos IoT por serem costumeiramente mais simples podem não conseguir executar as aplicações de machine learning.

Primeiramente vamos dividir os serviços entre dois tipos, aqueles que precisam de resultados em tempo real (usaremos aqui o critério de resultados em tempo menor do que 1s) e aqueles que podem receber os resultados em um outro momento ou que não precisam de resultado algum. Nessa solução iremos tratar apenas do primeiro grupo, tendo em vista que para o segundo grupo um simples processamento na *cloud* é o suficiente para uma solução aceitável. A abordagem proposta por alguns autores e reforçada no presente artigo é a utilização de conceitos e práticas de *edge computing* [4].

Edge computing é um paradigma de computação distribuída que traz o poder computacional e o armazenamento para perto de onde a solução é necessária, melhorando o tempo de resposta média das aplicações (por reduzir a latência) e diminuindo o uso de banda. Conceitos de *edge computing* já são utilizados em algumas redes de dispositivos IoT para diminuir o custo de processamento e o gasto energético, entretanto durante a pesquisa realizada para esse artigo foram encontradas poucas menções ao seu uso para suportar aplicações de machine learning [4].

Prevê-se que a utilização de um servidor local na rede dos dispositivos IoT possa reduzir drasticamente o custo e tempo de cada requisição feita pelos sensores, tendo em vista que além da latência mais baixa, podemos utilizar protocolos diferentes já que não iremos utilizar propriamente a *world wide web* e o HTTP, mas sim uma LAN que possui taxa de transferência superior. Além disso, ao utilizar essa abordagem o suporte dado ao usuário é melhorado, tendo em vista que normalmente ao utilizar *cloud computing* são usados parques de computadores terceirizados, enquanto ao utilizar *edge computing* as máquinas já estão próximas do local de uso.

Referências

- [1] M. Alsharif and D. B. Rawat, “Study of machine learning for cloud assisted iot security as a service,” *MDPI - Internet of Things*, 2021.

- [2] S. Mao, D. Bushmitch, S. Narayanan, and S. S. Panwar, "Agriculture iot market: Report," *Markets And Markets: Global Forecast to 2024*, 2020.
- [3] M. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, A. Ihsan, and M. Guizani, "A survey of machine and deep learning methods for internet of things (iot) security," *IEEE Communications Surveys and Tutorials*, vol. PP, 2020.
- [4] G. Vitali, M. Francia, M. Golfarelli, and M. Canavari, "Crop management with the iot: An interdisciplinary survey," *MDPI- Towards Artificially Intelligent Robotics for Agriculture – Current Developments and New Trends*, 2021.
- [5] K. H. Abdulkareem, M. A. Mohammed, A. Salim, M. Arif, O. Geman, D. Gupta, and A. Khanna, "Realizing an effective covid-19 diagnosis system based on machine learning and iot in smart hospital environment," *IEEE Internet of Things Journal*, pp. 1–1, 2021.
- [6] A. Sacco, M. Flocco, F. Esposito, and G. Machetto, "Crop management with the iot: An interdisciplinary survey," *Computer Communications*, vol. 160, pp. 769, 778, 2020.
- [7] H. Nguyen, "Machine learning model management using cloud-native technologies for iot," *Aalto University Master's Thesis Journal*, 2021.
- [8] V. e. a. HASSIJA, "A survey on iot security: application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019.
- [9] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. Hashem, "Attack and anomaly detection in iot sensors in iot sites using machine learning approaches," *Elsevier: Internet of Things*, vol. 7, 2019.
- [10] F. e. a. HUSSAIN, "Machine learning in iot security: current solutions and future challenges.," *IEEE Communications Surveys*, 2020.