



MACHINE LEARNING: CATALISADOR DA CIÊNCIA

NAS ÚLTIMAS DÉCADAS, É NOTÁVEL COMO A UTILIZAÇÃO DE MACHINE LEARNING IMPULSIONA AS DIVERSAS ÁREAS DO CONHECIMENTO CIENTÍFICO, PERMITINDO QUE PESQUISADORES CONCENTREM-SE EM ASPECTOS DE SUAS PRÓPRIAS ÁREAS DE ATUAÇÃO.

.....
por Héctor Allende-Cid

(Traduzido pelos Editores)
.....

Machine Learning é uma subárea da Inteligência Artificial que tem apresentado um crescimento enorme nas últimas décadas. Trata-se de algoritmos matemáticos, estatísticos e computacionais que são capazes de realizar um processo de inferência por meio de aprendizado baseado em exemplos. Há duas décadas, acreditava-se que os resultados alcançados por esses algoritmos seriam viáveis apenas nas mentes de autores de ficção científica.

Machine Learning é a área ideal para a automatização de processos, os quais podem ser "simples", como reconhecer padrões visuais, ou complexos, tais quais decisões de especialistas da área da saúde. Quando seres humanos lidam com problemas complexos, muitas vezes é impossível explicar o raciocínio que levou a tomar determinadas decisões. Por outro lado, é menos complexo realizarmos a coleta dos exemplos de decisões tomadas por seres humanos e usá-los como fonte para que o sistema aprenda a resolver o mesmo problema.

Com constante avanço científico e tecnológico em volta da área de ML, as demais áreas da ciência podem e/ou vão se beneficiar de técnicas de Machine Learning para acelerar e potencializar suas pesquisas. O ritmo acelerado do desenvolvimento de pesquisas nas áreas da Biologia, Medicina, Oceanografia, Linguística, etc. não permite que pesquisadores percam tempo realizando classificações manuais. A classificação é uma tarefa inerente em quase todas (quicá todas) as áreas das Ciências, automatizar essa tarefa é imprescindível para impulsionar o progresso científico [1].

Por exemplo, a Biologia Marinha e a Oceanografia realizam classificações e medidas de fitoplânctons em amostras de águas marinhas [2]. Nesse contexto, é de suma importância conhecer os diversos tipos de plâncton que se encontram nas amostras, uma vez que isso é um indicador da qualidade da água do mar. Essa informação é utilizada para definir políticas de pescas ou para saber o risco de potenciais problemas como a maré vermelha. A

contagem de plâncton é um tema de suma relevância para a investigação nesta área, pois essa tarefa geralmente é realizada em vários meses por especialistas de forma manual. Sem dúvidas, um processo automático de classificação e contagem de plânctons permitirá que os especialistas se concentrem em problemas mais complexos de sua área.

Machine Learning é a área ideal para a automatização de processos, os quais podem ser "simples", como reconhecer padrões visuais, ou complexos, tais quais decisões de especialistas da área da saúde.

Outro exemplo do potencial da Machine Learning é sua contribuição na área de Processamento de Linguagem Natural (PLN). É uma área interdisciplinar onde a inteligência artificial encontra a linguística para estudar as complexas relações entre os

computadores e as linguagens humanas. As linguagens naturais por si já são complexas por estarem em constante evolução e por serem uma ferramenta vital para estabelecer a comunicação com outros humanos e, assim, viver em sociedade. Há muitas tentativas de modelar as linguagens naturais, mas até o momento não há um modelo que seja capaz de refletir toda sua complexidade [3].

No Chile, estamos desenvolvendo uma pesquisa com profissionais da saúde e da linguística para a classificação automática de anamnese. A anamnese é o conjunto de dados que contém a história clínica dos pacientes. Atualmente, vários países têm esses dados registrados, mas a maioria sem a associação da sua respectiva enfermidade. Para um país, é de extrema importância conhecer as estatísticas das enfermidades mais relevantes de sua população, portanto classificar esses dados e identificá-los adequadamente é crucial para a elaboração e políticas pública eficientes. Utilizando uma grande quantidade de dados, as técnicas de ML permitem que esse processo de classificação e identificação seja feito de maneira automática e confiável.

As pesquisas apresentadas anteriormente são exemplos claros de como Machine Learning pode impulsionar o processo de investigação em diversas

áreas da ciência. Consequentemente, fazendo com que o conhecimento da humanidade avance em um ritmo muito maior. ●

Endereço com a versão original do artigo em espanhol:

<http://twixar.me/Fz0K>

Referências

1. Cynthia Rudin and Kiri L. Wagstaff. Machine learning for science and society. Machine Learning. 2014.
2. O. Py, H. Hong and S. Zhongzhi, "Plankton classification with deep convolutional neural networks. 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference. Chongqing, 2016.
3. Juliet Beni Edgcomb and Bonnie Zima. Machine Learning, Natural Language Processing, and the Electronic Health Record: Innovations in Mental Health Services Research. Psychiatric Services. 2019. MIT Press, url: <http://www.deeplearningbook.org>.



HÉCTOR ALLENDE-CID | É engenheiro de Informática e doutor em Engenharia de Informática pela Universidad Técnica Federico Santa María. Trabalha nas áreas de Machine Learning, Análise Preditiva, Computação Estatística. É professor da Escuela de Ingeniería Informática da Pontificia Universidad Católica de Valparaíso (Chile) e presidente da Associação Chilena de Reconhecimento de Padrões.