



### **A pesquisa na formação do cientista de dados**

Na atualidade, a pesquisa científica vê-se envolta pela vertente do “Data-Driven”, onde pesquisadores utilizam fortemente a computação como principal ferramenta analítica para obter conclusões de estudos em seus campos. Entretanto, é essencial saber que esse modo de pesquisa deve também ser acompanhado pelos métodos científicos tradicionais de formulação e verificação de hipóteses, pois estes constituem parte fundamental de um bom estudo.

*Machine Learning* é um termo cunhado no âmbito da ciência de dados e atualmente usado de forma ampla por pesquisadores em seus trabalhos. De maneira a assegurar resultados eficientes e mais rápidos, a comunidade científica abraçou vários modelos guiados unicamente por dados, deixando de lado em algumas ocasiões os métodos científicos tradicionais. Nesse sentido, o “fazer ciência” perde parte indispensável da sua constituição, visto que o uso e obtenção uníssonos de resultados por máquinas apenas nos fornece dados e informações, que, sem testes de hipóteses, sem a interdisciplinaridade e os questionamentos constantes da ciência, agregam pouco ou nada ao conhecimento humano.

Ademais, ao limitar-se à computação, o cientista pode perder acurácia, pelo fato de restringir melhores soluções para o problema. Logo, nesse cenário, fazer ciência atém-se à escrita de códigos e uso de algoritmos automatizados, onde o indivíduo vai gradualmente tornando-se um “pseudo-pesquisador”, já que não utiliza dos métodos científicos propriamente ditos e também não consegue entender, de forma profunda, seus resultados, nem mesmo como os obteve, pois não compreende completamente o que a máquina realiza.

Assim, é de suma importância que a pesquisa, em sua essência, contenha não somente modelos e ferramentas computacionais advindas da recente ciência de dados, mas também, ande em conjunto com os métodos científicos, a fim de tornar pesquisadores verdadeiros cientistas, e o conhecimento, menos dependente das máquinas.