

## Skills para data science.

Vamos aprender mais sobre as habilidades necessárias para aplicar o Machine Learning começando com uma pergunta. Você acredita que o usar Machine Learning requer muito tempo de treinamento em programação e algoritmos?

Talvez você tenha visto uma pesquisa de trabalho em Data Science que lista algoritmos como ferramentas. Considere os resultados no Kaggle, um site de competição de aprendizagem de máquinas onde as empresas carregam problemas não resolvidos e oferecem recompensas interessantes para os participantes.

Em problemas de alta complexidade, existem dois grupos principais de técnicas que normalmente dominam estas competições. Para problemas tabulares, nos quais o conjunto de dados pode ser claramente representado em uma tabela - por exemplo, prever uma coluna com o produto a ser recomendado a um usuário bancário com base em atributos de idade, renda, produtos que eles já contrataram. Nesses casos, geralmente domina alguma variante das técnicas de ensemble. Por exemplo, XG Boost.

Em problemas com dados não tabulares, imagens, áudio, alguma arquitetura de deep learning é normalmente a vencedora. Agora, realmente precisamos implementar esses modelos do zero?

Atualmente existem muitas soluções de machine learning como serviço. Em algumas etapas, pode-se colocar a entrada e a saída desejadas, definir as métricas e, em pouco tempo, obteremos um modelo altamente sofisticado, pronto para fazer previsões.

Podemos pensar em uma analogia com o desenvolvimento web. Para problemas que não são muito específicos, quase commodities, usamos uma plataforma pronta que evita todo o desenvolvimento, assim como usamos certos serviços para gerar uma página web sem escrever uma linha de código html. Agora, para problemas bem específicos, onde tanto a previsão como o desempenho do tempo de resposta são essenciais, precisaremos dos nossos próprios desenvolvedores.

Então qual é o passo-chave?

Em qualquer cenário, devemos fazer a pergunta certa, para avaliar as métricas ideais ao avaliar o modelo. Resumindo, se você possui dados de qualidade, organizados de uma forma que pode ajudar a classificar informações ou fazer previsões, você pode utilizar soluções prontas de Machine Learning, desde que você tenha alguém que lhe ajude a escolher os algoritmos apropriados.