

## **Resumo:** Designing Machine Learning Systems - Chip Huyen

No primeiro capítulo do livro “Designing Machine Learning Systems”, a autora apresenta a importância de compreender um sistema de Machine Learning (ML) como um todo. E a partir de uma visão geral, é explicitado como colocar um modelo ML em produção. Os algoritmos são apenas parte desse modelo. E antes de desenvolver um sistema de ML, faz-se necessário saber quando utilizar e quando não utilizar o ML. Posteriormente, são apresentados casos que o ML se faz necessário, e quando a solução para um determinado problema não precisa necessariamente de um ML. No entanto, mesmo que o ML não resolva o problema como um todo, indica-se dividir esse problema em partes e o ML pode ser útil para resolver parte desse problema.

O ML está em quase todos os aspectos de nossa vida. Como temos acesso à informação, como nos comunicamos, como trabalhamos. No entanto, há ainda muitos aspectos que o ML está em processo de desenvolvimento. É explicado os desafios enfrentados na implementação do ML que está em investigação diante dos ML que estão em produção. Do mesmo modo, são apresentadas as diferenças entre o ML e um software tradicional, incluindo um dado importante: que o ML gera padrões a partir dos dados de entrada e saída. Já o software tradicional precisa de padrões para gerar uma saída.

No segundo capítulo, Huyen apresenta a introdução ao design de ML, fundamentado de forma holística para que todos os componentes de um modelo de ML possam trabalhar juntos e satisfazer os objetivos e requisitos específicos do sistema. Para garantir que tudo está interligado, faz-se necessário compreender para quais objetivos o modelo de ML será orientado. Inicialmente é apresentado qual a intenção comercial e o contexto dos projetos de ML. É importante ressaltar que a dificuldade do trabalho pode mudar dependendo da forma como o problema é enquadrado, ou seja, ele pode se tornar mais difícil ou mais fácil. Huyen então afirma que para que um ML seja bem sucedido em uma empresa, o desempenho do ML deve estar alinhado ao desempenho geral da empresa.

Os requisitos para os sistemas de ML são apresentados e a autora relata que a maioria dos sistemas de ML devem ter quatro requisitos mais gerais: fiabilidade, escalabilidade, capacidade de manutenção e adaptabilidade. Ademais, um modelo de ML em produção é apresentado como um processo iterativo e também são descritos os passos desse processo, a saber: definição do escopo, engenharia de dados, desenvolvimento de modelos ML, implementação, monitoramento e análise comercial. Porém, esse processo se assemelha a um ciclo de idas e vindas.

Os progressos registrados na última década mostram que o sucesso de um sistema de ML depende em grande parte dos dados com que este modelo foi treinado. Logo, uma maior quantidade de dados nem sempre significa um melhor desempenho do ML. Porque dentre eles, podem existir dados desatualizados ou com informações incorretas.

No terceiro capítulo são apresentados os fundamentos da engenharia de dados. Neste capítulo a autora descreve as fontes de dados, a importância de escolher o formato correto para armazenar os dados, bem como são estruturados, para favorecer sua utilização no futuro. Os modelos de dados definem a forma como os dados são armazenados. Ainda, são apresentados os três principais modelos, a saber: relacional, documento e gráfico. Portanto, estes capítulos tem como objetivo fornecer uma compreensão da estrutura do ML para desenvolver uma solução que melhor resolva um problema com um sistema de ML.