

## A Importância do Machine Learning e suas Aplicação em Produção Industrial

O texto inicia apresentando uma visão geral sobre os requisitos e desafios envolvidos na implementação de sistemas de aprendizado de máquina (ML) no mundo real, destacando-se a importância de compreender as diferenças entre ML em ambiente de pesquisa e sua aplicação prática em escala. Uma vez que, existem disparidades entre ML em ambiente acadêmico e sua implementação em produção, é relevante observar aspectos como envolvimento das partes interessadas, prioridade computacional, propriedades dos dados, questões de ética e interpretabilidade. Essa distinção é fundamental para os profissionais que transitam da academia para o desenvolvimento de soluções ML em ambiente real.

Baseado em casos práticos, evidencia-se que, a maioria das aplicações de uso ocorre no contexto empresarial. Além disso, são abordados as circunstâncias em que as soluções de ML devem se encaixar com base nos dados dessa forma, aprendendo padrões complexos e posteriormente realizando previsões com novos dados. Uma ênfase especial é dada à complexidade dos sistemas de ML, os quais, englobam uma variedade de componentes além dos algoritmos de ML, como: infraestrutura, pilha de dados, implantação, monitoramento e manutenção.

A discussão avança destacando que, o desenvolvimento de sistemas de ML depende crucialmente da seleção e avaliação adequadas dos modelos e objetivos de negócios, uma vez que, a meta final de qualquer projeto dentro de uma empresa é de modo geral, aumentar os lucros, direta ou indiretamente. É abordada uma ampla gama de modelos de ML, desde os mais simples até os mais avançados, sempre considerando quatro aspectos: confiabilidade, escalabilidade, capacidade de manutenção e adaptabilidade. Esses requisitos apresentam características, aplicações e desafios associados, de tal forma que muitas empresas criam seus próprios parâmetros para mapear métricas de negócios em métricas de ML.

É importante destacar a escolha de modelo mais adequado para o problema em questão, levando em consideração fatores como complexidade do problema, tamanho e natureza dos dados, e requisitos de interpretabilidade. Estratégias para avaliação e comparação de modelos, como validação cruzada e análise de erros são essenciais na tomada de decisões, uma vez que todo o sistema deve continuar a executar suas funções de forma correta. Esse sistema precisará ser continuamente monitorado e atualizado, ou seja, se faz necessário de uma boa manutenção e para isso é de boa prática dissociar os múltiplos objetivos do projeto.

Por fim, dar-se foco na função dos dados, a escolha de formatos de armazenamento adequados, modelos de dados principais e mecanismos de armazenamento e processamento. É importante entender a distinção entre dados estruturados e não estruturados, modelos de dados (relacional, documento e gráfico). Outro ponto de grande relevância é a tráfego de dados entre processos, que pode ser processada em lote ou em fluxo de tempo real, onde cada um tem suas vantagens técnicas a depender da aplicação.

A importância dos dados se dá pela ascensão do ML nos últimos anos conectada a big data. A grande quantidade de dados é sinônimo de diversidade de processos e serviços, logo é muito mais interessante manter esses dados armazenados de forma bruta e conforme a necessidade essas informações são consumidas. Existe um ponto de atenção quanto a entrada de dados pelo usuário, pois exigem uma verificação e processamento mais específicos, para que se mantenha um padrão. Padrão é a palavra-chave nos processos de armazenamento e processamento de dados, pois temos uma quantidade bem variada de dados como texto, vídeo, áudio, entre outros.