

Atividade 1

1. Machine learning é um sistema capaz de aprender com os dados, invés de serem programados diretamente para cada tarefa.
2. Conjunto de treinamento é o conjunto de dados usado para ensinar o modelo, a partir disso, o modelo irá analisar esses dados e ajustar seus parâmetros internos para aprender padrões e relações; Conjunto de Validação é usado para ajustar e avaliar o modelo durante o treinamento. Seus objetivos principais são avaliar o desempenho do modelo e ajustar hiperparâmetros; já o conjunto de teste é utilizado somente no final, após todo o treinamento e ajuste do modelo. Esse conjunto irá avaliar a capacidade de generalização do modelo em dados completamente novos.
3. Os modos mais eficazes para lidar com dados ausentes seriam: remover os dados ausentes, em caso de os valores serem pequenos ou o valor não é essencial para o modelo; preencher os dados ausentes, usando métricas como média, mediana, moda e, até mesmo, regressão.
4. A matriz de confusão é uma tabela usada para avaliar o desempenho de modelos de classificação em ML. Essa matriz mostra quantas previsões o modelo acertou e quantas errou, fazendo uma comparação com os valores previstos com os valores reais.
5. A aplicação de algoritmos de ML é promissora em diversas áreas, com suas diversas nuances e destaques. Levando em consideração, por exemplo, a área de manufatura, os principais ganhos podem ser: otimização da manutenção preditiva, controle de qualidade automatizado e otimização de processos industriais, por meio da utilização de sensores e produção em tempo real. Essas implementações, dentre outras, gerariam mais eficiência, menos paradas e acidentes, e um menor custo operacional.