

ATIVIDADE 01 – CURSO MACHINE LEARNING 2024.1

Aluno: Rafael Kashirajima

1. Explique, com suas palavras, o que é machine learning?

É uma área da inteligência artificial que desenvolve algoritmos e modelos que aprendem a partir de dados. Tem a finalidade de apoiar tomadas de decisões e fazer previsões.

2. Explique o conceito de conjunto de treinamento, conjunto de validação e conjunto de teste em machine learning.

A base de dados pode ser dividida em dados para treino, dados para teste e dados para validação.

Os dados de treino são utilizados como entrada para o algoritmo adquirir o aprendizado.

Os dados de teste são utilizados após o aprendizado a fim de testar o modelo.

Os dados de validação são usados para avaliar o desempenho do modelo nas fases de treinamento e testes e ajuda a ajustar os parâmetros.

3. Explique como você lidaria com dados ausentes em um conjunto de dados de treinamento.

Depende da situação e das regras de negócios.

Exemplos de abordagem:

- preencher os dados usando valores estatísticos.
- calcular
- Buscar o dado
- ignorar, caso o dado não seja necessário.

4. O que é uma matriz de confusão e como ela é usada para avaliar o desempenho de um modelo preditivo?

Uma tabela com os dados resultantes da avaliação da fase de teste de um algoritmo de classificação.

Na tabela, serão especificados a quantidade de verdadeiros(positivos e negativos) e falsos(positivos e negativos).

A matriz pode ser usada para calcular métricas de desempenho.

5. Em quais áreas (tais como construção civil, agricultura, saúde, manufatura, entre outras) você acha mais interessante aplicar algoritmos de machine learning?

Acho interessante o uso de machine learning para auxiliar a tomada de decisões nos negócios, através de previsões e identificação de padrões.