



CURSO: <BÁSICO EM MACHINE LEARNING>

- **Atividade 01 (ATIV-01)**

- Tipo: Diagnóstica.
- Tema: Conceitos básicos de machine learning.
- Conteúdo: Machine learning.
- Participantes: Individual.
- Avaliação do aluno.
 - Objetivo: Avaliar o conhecimento do aluno sobre conceitos básicos de machine learning.
 - Nota: Sem valor numérico (apenas um guia para o processo educacional).
 - Critérios avaliados: Respostas com coerência, coesão e com exemplos.
- Informações complementares: A atividade é composta por 5 questões dissertativas. Os resultados são utilizados para divisão de grupos da atividade 04;
- **AO CONCLUIR A ATIVIDADE: ENVIAR APENAS O LINK DO REPOSITÓRIO GITHUB (ESPECIFICAR A BRANCH) PÚBLICO.**

1. Explique, com suas palavras, o que é machine learning?
Machine Learning (Aprendizado de Máquina) é uma maneira de ensinar computadores a aprender padrões a partir dos dados que são fornecidos e tomem decisões ou façam previsões sem precisar ser programados com regras específicas para cada situação.
2. Explique o conceito de conjunto de treinamento, conjunto de validação e conjunto de teste em machine learning.

Conjunto de Treinamento:

É o conjunto usado para ensinar o modelo. O algoritmo aprende os padrões e relações presentes nos dados durante essa fase.

Conjunto de Validação:

É o conjunto usado para validar o treinamento do modelo enquanto ele ainda está sendo desenvolvido. Esse conjunto ajuda a evitar que ele memorize apenas os dados de treinamento.

Conjunto de Teste:

Este conjunto é usado após o treinamento do modelo para avaliar seu desempenho. Este conjunto contém dados que o modelo nunca viu antes, permitindo verificar sua capacidade de generalização para situações reais.

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

Já trabalhei com conjuntos de dados que possuíam valores faltantes e utilizei algumas técnicas para lidar com esse problema. Para tratar dados ausentes em um conjunto de dados de treinamento, primeiro é necessário analisar a quantidade e o tipo de dados faltantes para escolher a melhor forma de tratamento.

Uma das opções é remover registros ou colunas que possuem valores ausentes, principalmente quando a quantidade de dados faltantes é pequena e não compromete a análise. Outra alternativa é substituir os valores ausentes utilizando técnicas de imputação, como preencher com média, mediana ou moda, dependendo do tipo da variável.

Também é possível utilizar modelos estatísticos ou algoritmos de machine learning para estimar os valores ausentes com base em outros dados disponíveis.

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

A matriz de confusão é uma ferramenta de visualização e avaliação que é fundamental em ML. A matriz apresenta quatro resultados principais: verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo.

É usada especificamente para problemas de classificação. Ela permite ver não apenas o erro total do modelo, mas onde ele está errando.



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

Os algoritmos de Machine Learning podem ser aplicados em diversas áreas. A área da saúde é uma das mais interessantes, pois permite auxiliar no diagnóstico de doenças, na análise de exames médicos e na previsão de surtos, contribuindo para tratamentos mais rápidos e precisos para os pacientes.

Na construção civil, pode ser utilizado para reduzir custos ao apontar possíveis falhas estruturais, ajudando no planejamento e na redução de gastos com mão de obra e manutenção.

Na área agrícola, pode auxiliar na previsão de safras, no monitoramento do clima, na identificação de doenças e no controle de pragas nas lavouras. Entre essas áreas, a saúde se destaca pelo impacto direto na qualidade de vida das pessoas, mas todas apresentam grande potencial de crescimento com o uso de Machine Learning.