**CURSO: Básico Machine Learning - Avanti**

**Atividade 01**

**Aluna: Madelu Lopes Da Silva**

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

Machine learning é uma área que está ramificada à Inteligência Artificial, ou seja, pertence a ela. Em outras palavras, Machine learning, ou aprendizado de máquina é a maneira como as máquinas imitam a forma como os seres humanos aprendem. Isso se dá através de dados, onde estes sistemas são treinados usando dados e algoritmos para garantir um bom desempenho no seu aprendizado.

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

* **Conjunto de Treinamento:** São subconjunto de dados utilizado para treinar o modelo. Os dados são retirados de um dataset e utilizado para que o modelo aprenda padrões e relações.
* **Conjunto de Validação:** Conjunto de dados para validação frequente do modelo.
* **Conjunto de Teste**: Conjunto de dados utilizados para avaliar o desempenho do modelo.

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

Retiraria os valores se percebesse que eles estão em uma coluna que não tem necessidade de estar no dataset, ou seja, se forem dados que possam ser removidos e sua falta não fara diferença. Caso contrário, posso substituir os valores pela média, mediana ou moda se assim for viável, também posso colocar valores através de padrões numéricos, ou mesmo de nome isso quando os dados mostram estes padrões.

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

É uma tabela especifica que tem a função de permitir a visualização da performance do algoritmo de classificação. Os valores previstos pelo modelo são comparados com os valores reais dos dados, permitindo visualizar os modelos está acertando e onde está errando.

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

Saúde, educação, sistemas de vigilância em vídeo, bots de serviço ao cliente, segurança em T.I, análise de streaming de dados (especialmente em música), processamento de linguagem natural e mecanismo de busca.