**Atividade 01 (ATIV-01)**

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

Nos últimos dois anos eu tenho escutado muito o termo aprendizado de máquina, eu trabalho na indústria e muitos fornecedores tem nos oferecido soluções das mais diversas áreas, como sistemas que permitem avaliar e prever quebras de equipamentos, como câmeras com sistema integrado com visão computacional, automatização de equipamentos por meio de visão computacional, sistemas de automatização de controle de estoque inteligentes e por ai vai, tenho percebido bastante essa mudança no mercado por meio das IAs, comecei a estudar desde o ano passado sobre IoT e indústria 4.0, me interessei bastante sobre esse mundo, ai percebi que é necessário estudar sobre programação para desenvolver esses sistemas, **machine learning** no meu conceito engloba bastante dessas aplicações citadas ai em cima, vão desde sistemas de recomendações em [streamers](https://www.google.com/search?sca_esv=795b186f858532d3&rlz=1C1ONGR_pt-PTBR1087BR1087&sxsrf=AE3TifM2pLbXh5NEI6G2bv4nJAb-nOIzXA:1749420396712&q=streamers&spell=1&sa=X&ved=2ahUKEwikhrzq6uKNAxVJIrkGHcK_IMcQkeECKAB6BAgKEAE), redes sociais e etc, como as mais diversas áreas de saúde, finanças, agro, industrias, interagimos como interagimos machine learning o tempo todo, é uma tecnologia muito importante nos dias de hoje, muito utilizada para identificar padrões, fazer previsões com base no que aprendeu com os dados fornecidos, uma área que tenho muito curiosidade e vontade de aprender é visão computacional, por isso estou fazendo este curso, mas tenho pouco conhecimento e vivência em programação.

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

- No treinamento é onde os dados que foram alimentados do modelo, serão utilizados para treinar o programa, geralmente utilizam uma média de 70% dos dados para treinar o modelo, esses valores não são fixos, variam de acordo com a necessidade e ajustes, aqui o programa tentará encontrar padrões, se o programa errar durante o treinamento, ele ajustará os seus parâmetros para melhorar o seu acerto da próxima vez e assim por diante.

- Na validação com base nos resultados, é onde é avaliado o desempenho de acertos e erros do treinamento do modelo, verifica também se o programa é “viciado nos dados”, não lembro o nome do termo para isso.

- Na etapa de testes, é onde é testado o modelo com dados ainda não visto pelo o programa, se o desempenho com dados novos for bom, indica que o modelo é bom e pode seguir para a próxima etapa de implementação, se o desempenho dor ruim, terá que voltar para o início do modelo e realizar novos ajustes, seja por alimentar o modelo com novos dados ou ajustes nos parâmetros.

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

Essa etapa de limpeza dos dados, geralmente se não for um dado que irá impactar no desempenho do modelo, o indicado é deletar ele da base de dados, mas dependendo dos dados, também tem a opção de fazer uma média, mediana ou moda para substituir os dados faltantes.

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

Eu já ouvir falar uma vez sobre esse termo, não me recordo direito, talvez tenha algo a ver sobre avaliar se o modelo está viciado com os dados e perca a capacidade de classificar e avaliar dados novos nunca visto? Posso estar errado, mas estou respondendo essas questões sem pesquisar na internet, estou fazendo outro curso de ciência de dados, e tive essa experiência em introdução a redes neurais.

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

Na saúde é muito importante para auxiliar os médicos em diagnósticos de doenças, também acho bastante importante essa parte de visão computacional, voltado tanto para segurança como para identificar objetos, as possibilidades são imensas para todas as áreas.