

Relatório 6 -Vídeo: O Que É Ciência De Dados (I)

Higor Miller Grassi

No primeiro vídeo do nerdologia é falado sobre ciência de dados, que vai além da simples análise de informações, ele acaba envolvendo o uso inteligente de dados para extrair insights que muitas vezes não estão diretamente visíveis, muitas vezes não se tratando apenas de observar o que os dados mostram, mas de interpretá-los para entender o que está por trás das informações e prever possíveis cenários futuros, enxergando “fora da caixa”, ou seja, enquanto os dados brutos nos contam o que já aconteceu, a ciência de dados permite usar esse conhecimento para tomar decisões informadas e antecipar eventos que ainda estão por vir, onde matematicamente falando, a estatística desempenha um papel fundamental, já que é a ferramenta matemática que oferece a base para essas análises.

Uma ferramenta poderosa que complementa a estatística dentro da ciência de dados é o machine learning (aprendizado de máquina), se tornando importante quando os dados que precisam ser analisados não estão em formatos tradicionais, como números ou tabelas estruturadas, como texto, vídeos e imagens de satélites, nesses casos, algoritmos de machine learning são capazes de processar e extrair padrões desses dados complexos de uma maneira que os métodos tradicionais de análise estatística não conseguem, dito isso, os dados analisados precisam de qualidade, pois com dados ruins podem levar a conclusões equivocadas e a modelos que falham ao serem aplicados em cenários reais.

Já no segundo vídeo, se trata de uma parte mais profunda sobre explorar o poder dos dados para criar o maior impacto possível para uma empresa, seja por meio de produtos de dados ou melhorias de produtos existentes, onde ao aplicar técnicas de ciência da computação e mineração de dados, graças ao William S. Cleveland que sugeriu integrar métodos de ciência da computação ao campo da estatística, é possível transformar grandes volumes de dados em insights valiosos, possibilitando uma tomada de decisão mais precisa e informada.

THE DATA SCIENCE HIERARCHY OF NEEDS

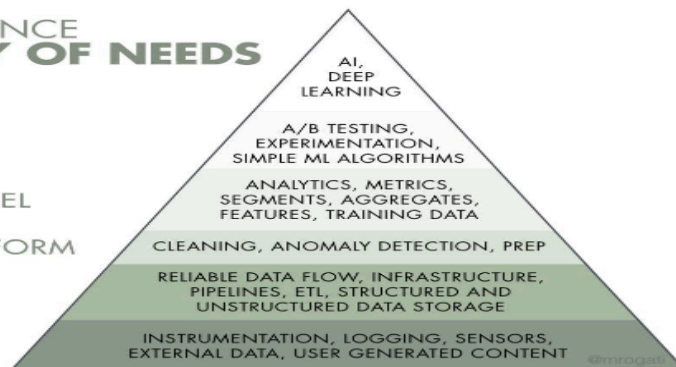
LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT



A hierarquia das necessidades em ciência de dados, conforme ilustrada na imagem, inicia com a coleta e o armazenamento de dados, avança para a transformação e análise, e finalmente atinge a otimização e o aprendizado profundo (deep learning) que dito anteriormente no curso, é um ramo da inteligência artificial que utiliza redes neurais artificiais para simular o funcionamento do cérebro humano e resolver problemas complexos. Sendo falado também que as diferentes empresas buscam maneiras distintas dentro desta hierarquia, para aprimorar e contratar profissionais hábeis ao serviço.

Conclusão: Mostra a importância da ciência de dados, onde nos apresenta que a coleta de dados serem efetuadas corretamente para poder treinar as máquinas de forma que atue no mundo real. Diz também que, os cientistas de dados, tornam-se resolvidores de problemas e estrategistas, tendo que lidar com o pensamento de o que aconteceu no passado, permanece no presente e provavelmente se perpetuar no futuro, para assim formular a melhor estratégia para resolver este impasse.

Referências:

▶ [O que é ciência de dados | Nerdologia Tech](#)

▶ [What REALLY is Data Science? Told by a Data Scientist](#)