

Aula 01: Introdução à Ciência de Dados e a Pergunta de Negócio

Professor: Paulo Rogério Pires Manseira

Alunos: Maruan Biasi El Achkar e Ricardo Falcão Schlieper

Parte 1

- **O que é ciência de dados?**

É o processo de analisar grandes volumes de dados usando estatística, programação e conhecimento do negócio para gerar insights e tomar decisões, tem um fluxo de trabalho.

- **Quem pratica e suas diferenças?**

Cientista de Dados: Desenvolve modelos e algoritmos para prever comportamentos e identificar padrões em dados complexos.

Analista de Dados: Explora e interpreta dados, gerando relatórios e insights para apoiar decisões.

Engenheiro de Dados: Constrói, organiza e mantém a infraestrutura que permite o uso eficiente dos dados.

Parte 2

- **Como Funciona na Prática? O Ciclo de Vida dos Dados**

É estruturado em 6 fases: Entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação e implantação. O ciclo de vida serve como um guia para projetos de ciência de dados, facilitando a execução de um projeto de sucesso, podendo ser adaptada dependendo dos requisitos do projeto

Parte 3

- **Cenário**

O cenário escolhido é a fraude em transações de cartão de crédito, um problema crescente para bancos e clientes devido ao aumento das compras dessa modalidade e do avanço nas técnicas de golpe. O dataset reúne informações como distância da compra em relação à residência e à compra anterior, se a loja já foi utilizada, método de pagamento, se foi inserido ou aproximação, se a senha foi utilizada, como o valor da compra se compara a média de compras, etc. A partir desses dados, vamos identificar quais fatores estão mais relacionados com fraudes

- **Pergunta de negócio**

Quais fatores indicam que uma compra com cartão de crédito é fraudulenta?