Aluno: Marco Leone Merini

Data: 05/08/2025

Professor: Paulo Manseira

**Parte 1: O Que é Ciência de Dados e Quem a Pratica?**

**a) O que é Ciência de Dados em suas próprias palavras?**  
Ciência de Dados é uma área que combina estatística, programação e conhecimento de negócio para extrair insights de dados. Seu objetivo é resolver problemas e apoiar a tomada de decisões com base em evidências extraídas de grandes volumes de dados.

**b) Quais as diferenças entre Cientista de Dados, Analista de Dados e Engenheiro de Dados?**

* **Cientista de Dados**: Cria modelos preditivos, realiza análises complexas e propõe soluções baseadas em dados.
* **Analista de Dados**: Foca na análise descritiva e diagnóstica dos dados, gerando relatórios e dashboards.
* **Engenheiro de Dados**: Constrói e mantém a infraestrutura e os pipelines para coleta, armazenamento e processamento dos dados.

**Parte 2: Como Funciona na Prática? O Ciclo de Vida dos Dados**

O ciclo de vida de um projeto de ciência de dados normalmente segue o modelo **CRISP-DM**, que é uma metodologia estruturada e amplamente utilizada na área. Suas etapas são:

1. **Entendimento do Negócio:**  
   Compreensão do problema que a empresa enfrenta e definição dos objetivos do projeto de dados. Exemplo: aumentar a retenção de usuários em um aplicativo.
2. **Entendimento dos Dados:**  
   Coleta e exploração dos dados disponíveis, avaliação da qualidade, identificação de dados relevantes e possíveis limitações.
3. **Preparação dos Dados:**  
   Etapa que envolve limpeza, transformação, integração e seleção dos dados necessários para análise. Geralmente, essa é a etapa mais trabalhosa do projeto.
4. **Modelagem:**  
   Aplicação de algoritmos e técnicas estatísticas/matemáticas para construir modelos preditivos ou descritivos. Inclui escolha dos métodos e ajuste de parâmetros.
5. **Avaliação:**  
   Verifica se os modelos atendem aos objetivos do negócio. São usados indicadores como acurácia, precisão, recall e validação cruzada.
6. **Implantação:**  
   Colocação do modelo em produção ou geração de relatórios e recomendações. A entrega final pode ser um sistema automatizado, um dashboard ou uma apresentação para a diretoria.

**Parte 3**

**Área escolhida:** Saúde e Bem-estar – App de Fitness

**Cenário**

Uma empresa desenvolvedora de aplicativos de bem-estar deseja melhorar a retenção de usuários que utilizam dispositivos como o Fitbit. Dados mostram que muitos usuários abandonam o uso do aplicativo após poucas semanas, apesar de terem acesso a recursos de monitoramento de sono, passos, calorias e atividade física.

Com base no histórico de atividades registrado pelos dispositivos Fitbit, o objetivo é descobrir o que motiva os usuários a continuarem utilizando o app e o que pode prever sua desistência. A empresa quer desenvolver estratégias personalizadas de engajamento, como alertas, gamificação ou recomendações inteligentes.

**Pergunta de Negócio (boa):**

**"Quais métricas de uso (ex: frequência de passos, calorias queimadas, tempo de sono ou nível de atividade física) mais influenciam a continuidade do uso do aplicativo após as primeiras 4 semanas?"**