**Definição do Problema**

**Com base no dataset “vgsales.csv”**

**1. Entendimento do Problema**

Antes de qualquer coisa, é essencial compreender o contexto do problema. Para isso, precisamos responder a perguntas-chave:

* **Qual é o problema exato a ser resolvido?**
* **Quais são os objetivos e necessidades do negócio?**
* **Quais perguntas específicas precisam ser respondidas?**
* **Quais métricas de sucesso serão usadas para avaliar a solução?**

**2. Ferramentas para Definição do Problema**

**Revisão da Literatura**

A revisão da literatura científica é crucial para entender o estado da arte e as abordagens já existentes. Pesquisamos artigos acadêmicos, publicações de conferências e patentes para encontrar métodos e algoritmos que foram utilizados para resolver problemas semelhantes.

**Estudos de Caso Existentes**

Analisar estudos de caso permite identificar práticas bem-sucedidas e comuns em indústrias semelhantes. Estudos de caso também ajudam a entender como diferentes problemas foram abordados e quais soluções se mostraram eficazes.

**Práticas Recomendadas**

As práticas recomendadas em ciência de dados e aprendizado de máquina fornecem diretrizes úteis para a definição de problemas. Isso inclui frameworks como CRISP-DM (Cross Industry Standard Process for Data Mining), que fornece uma abordagem estruturada para a definição e solução de problemas de dados.

**Exemplo Prático: Previsão de Demanda em Varejo**

**Entendimento do Problema**

**Problema:** Uma empresa de varejo quer prever a demanda de produtos para otimizar seu estoque e reduzir custos de armazenamento.

**Necessidades do Negócio:** A empresa precisa de previsões precisas para evitar excesso ou falta de estoque, melhorar a eficiência operacional e aumentar a satisfação do cliente.

**Perguntas-chave:**

* Quais fatores influenciam a demanda dos produtos?
* Como esses fatores variam ao longo do tempo?
* Qual é a precisão esperada das previsões?

**Métricas de Sucesso:**

* Redução de custos de armazenamento.
* Aumento da precisão das previsões.
* Melhora na disponibilidade dos produtos.

**Revisão da Literatura**

Pesquisas anteriores mostram que fatores como sazonalidade, campanhas de marketing e eventos especiais têm um impacto significativo na demanda de produtos no varejo.

**Estudos de Caso**

Um estudo de caso de uma grande rede de supermercados utilizou modelos de aprendizado de máquina, como Regressão Linear e Random Forest, para prever a demanda e conseguiu reduzir os custos de armazenamento em 15%.

**Práticas Recomendadas**

**CRISP-DM:** Seguir as etapas do CRISP-DM pode ajudar a estruturar a solução:

1. **Compreensão do Negócio:** Entender os objetivos e restrições.
2. **Compreensão dos Dados:** Coletar e explorar os dados disponíveis.
3. **Preparação dos Dados:** Limpar e transformar os dados para análise.
4. **Modelagem:** Selecionar e aplicar modelos de previsão.
5. **Avaliação:** Avaliar a precisão dos modelos.
6. **Implementação:** Implementar o modelo selecionado em produção.