



***Big Data Real-Time Analytics Com Python e Spark 3.0***

# **Big Data Real-Time Analytics Com Python e Spark Versão 3.0**

## **Quando Usar Label Encoding e One Hot Encoding?**

Essa pergunta geralmente depende do seu conjunto de dados e do modelo que você deseja aplicar. Mas ainda assim, há uma regra geral a ser observada antes de escolher a técnica de codificação correta para o seu modelo:

**Aplicamos Label Encoding quando:**

- A variável categórica é ordinal (como por exemplo variável que indica escolaridade: ensino fundamental, ensino médio, etc...). Há uma ordem nas categorias da variável.

**Aplicamos One-Hot Encoding quando:**

- A variável categórica não é ordinal, ou seja, a variável é nominal (como por exemplo variável país). Nesse caso não há uma ordem nas categorias da variável.

Essas são apenas regras gerais ficando sob a responsabilidade do Cientista de Dados a decisão da técnica a ser usada. Lembre-se de sempre justificar as suas decisões, demonstrando assim o seu conhecimento.