



Microsoft Power BI Para Business Intelligence e Data Science

Microsoft Power BI Para Business Intelligence e Data Science

O Que é Engenharia de Atributos?

A Engenharia de Atributos (também conhecida como Feature Engineering, em inglês) é o processo de criação, seleção e transformação de atributos (variáveis ou características) em um conjunto de dados. Essa técnica é usada para melhorar a performance dos modelos de aprendizado de máquina, além de tornar os dados mais informativos e relevantes para a análise.

A Engenharia de Atributos envolve várias etapas e técnicas, incluindo:

Seleção de atributos: Identificar e selecionar os atributos mais importantes e relevantes para o problema em questão. Isso pode envolver a remoção de atributos com alta correlação, redundantes ou irrelevantes.

Transformação de atributos: Modificar ou transformar os atributos existentes para torná-los mais úteis para análise ou criação de modelos de Machine Learning. Isso pode incluir normalização, padronização, discretização, aplicação de funções logarítmicas ou exponenciais, entre outras.

Criação de novos atributos: Criar novos atributos a partir dos existentes, combinando-os ou aplicando funções matemáticas. Isso pode envolver a criação de atributos polinomiais, interações entre atributos, atributos derivados de funções matemáticas, entre outros.

Tratamento de valores ausentes: Lidar com valores ausentes (missing values) no conjunto de dados, substituindo-os por medidas estatísticas, como média, mediana ou moda, ou utilizando técnicas mais avançadas, como a interpolação ou a imputação por modelos de aprendizado de máquina.

Codificação de variáveis categóricas: Converter variáveis categóricas em representações numéricas que possam ser utilizadas pelos modelos de aprendizado de máquina. Isso pode incluir a aplicação de técnicas como one-hot encoding, ordinal encoding ou label encoding.

A Engenharia de Atributos é uma etapa crucial no processo de construção de modelos de aprendizado de máquina, uma vez que a qualidade e a relevância dos atributos têm um impacto significativo na performance do modelo. A escolha e a criação de atributos apropriados podem levar a modelos mais precisos, eficientes e generalizáveis.

O tema é amplamente estudado em diversos cursos aqui na DSA através de diversas técnicas e ferramentas, e vamos trazer agora para você uma breve demonstração de como realizar Engenharia de Atributos com Linguagem M no Power BI.