



Big Data Real-Time Analytics Com Python e Spark 3.0

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

A Importância da Engenharia de Atributos

A Engenharia de Atributos é a 'arte' de formular recursos úteis a partir de dados existentes seguindo o destino a ser aprendido e o modelo de aprendizado de máquina a ser usado.

Envolve transformar dados em formas que melhor se relacionem com o alvo a ser aprendido. Quando bem feita, a Engenharia de Atributos pode aumentar o valor dos dados existentes e melhorar o desempenho dos modelos de Machine Learning. Por outro lado, usar recursos ruins pode exigir que você construa modelos muito mais complexos para atingir o mesmo nível de desempenho.

Essa é a razão pela qual a Engenharia de Atributos encontrou seu lugar como uma etapa indispensável no pipeline de aprendizado de máquina. No entanto, quando se trata de aplicar esse conceito, não há um método rígido e rápido ou estrutura teórica e é por isso que esta é uma das etapas mais complexas dentro do processo de Ciência de Dados, sendo uma das mais importantes para uma análise mais precisa e modelos de Machine Learning menos propensos a erros.