Machine Learning and Data Mining

(Mitchell, 1999)

Mitchell, T. M. (1999). Machine learning and data mining. *Communications of the ACM*, *42*(11), 30–46. https://doi.org/10.1145/319382.319388

Durante a última década, muitas organizações começaram a recolher regularmente enormes volumes de dados históricos descrevendo as suas operações, produtos, e clientes. Ao mesmo tempo, cientistas e engenheiros em muitos campos têm vindo a capturar dados experimentais cada vez mais complexos conjuntos, tais como gigabytes de dados de ressonância magnética funcional (MRI) descrevendo a actividade cerebral nos seres humanos. O área de mineração de dados aborda a questão de como melhor utilizar estes dados históricos para descobrir as regularidades gerais e melhorar o processo de tomada de decisões.

O crescente interesse na prospecção de dados, ou na utilização de dados históricos para descobrir regularidades e melhorar decisões futuras, decorre da confluência de várias tendências recentes: a queda do custo dos grandes dispositivos de armazenamento de dados e a crescente facilidade de recolha de dados através de redes; o desenvolvimento de algoritmos robustos e eficientes de aprendizagem de máquinas para processar estes dados; e a queda do custo da potência computacional, permitindo a utilização de métodos computacionalmente intensivos para a análise de dados. O campo da mineração de dados, por vezes denominado "descoberta de conhecimento a partir de bases de dados", "análise avançada de dados", e aprendizagem de máquinas, já produziu aplicações práticas em áreas como a análise de resultados médicos, detecção de fraudes com cartões de crédito, previsão do comportamento de compra dos clientes, previsão dos interesses pessoais dos utilizadores da Web, e optimização dos processos de fabrico. Também levou a um conjunto de questões científicas fascinantes sobre como os computadores podem aprender automaticamente com a experiência passada.