

Big Data e Data Science - uma visão estatística

Amílcar Oliveira

(Professor Auxiliar, amilcar.oliveira@uab.pt)

Universidade Aberta, Departamento de Ciências e Tecnologia e CEAUL,
Portugal

Resumo

Big Data (BD) tem potencial para avaliar e melhorar o processo de tomada de decisões e nos últimos anos atraiu forte interesse de académicos e profissionais. BD tornou-se o novo paradigma no mundo da recolha e análise de dados sendo geralmente caracterizado por cinco dimensões: volume (número de observações, atributos e relações), variedade (diversidade de fontes de dados, formatos, mídia e conteúdo), velocidade (velocidade de produção e mudança dos dados), veracidade (qualidade, origem e confiabilidade dos dados) e valor. *Big Data Analytics* (BDA) está a tornar-se cada vez mais um gerador de tendências que muitas organizações adotam com o propósito de criar informações valiosas de BD. As áreas de Ciências da Computação, Engenharia e Estatística, desempenham um papel fundamental na análise do BD, cada uma com a sua especificidade, mas todas igualmente importantes. *Data Science* (DS) é uma área interdisciplinar dedicada ao estudo e à análise de dados, que visa a extração de conhecimento para possíveis tomadas de decisão. DS combina entre outras, as áreas de Ciências da Computação, Matemática e Estatística, numa perspetiva de atuação num domínio do qual são provenientes os dados, tendo a capacidade de transformar grandes quantidades de dados em *insights* de negócios, e com isso, auxiliar instituições e empresas na tomada de decisões para atingir melhores resultados. Neste seminário pretende-se apresentar uma visão geral e atual sobre esta temática e discutir alguns dos desafios da Estatística nesta nova era do *Data Science* e do *Big Data*.