

DATA SCIENCE

VII Encontro de Bolsistas de Apoio a Projetos da Graduação

Maria de Lourdes Maia Silva, Julio Francisco Barros Neto

A pesquisa consiste em uma divulgação científica e tem como objetivo introduzir Data Science e apresentar as vantagens que a tecnologia oferece ao mercado, sendo abordadas as áreas de conhecimento, áreas de atuação e formação de um cientista de dados. Devido a era do Big Data, a ciência de dados está se tornando um campo muito promissor para explorar e processar grandes volumes de dados gerados por diferentes fontes. Um exemplo da evolução desse processamento de dados é o uso de softwares em micro e pequenas empresas. O interesse em banco de dados surgiu pois permitem o acesso direto às informações geradas pela interpretação do conjunto de dados, que, antigamente, eram arquivados manualmente, o que levava horas. No estudo do Big Data, existem dois conceitos muito importantes, o desenvolvimento do processo de análise, que consiste na evolução dos dados à sabedoria, os dados representam os valores a serem analisados, após a interpretação e contextualização dos dados, temos acesso à informação, após a modelagem e aplicação de padrões, chegamos ao conhecimento e à otimização dos negócios e tomadas de decisões que representam a sabedoria. Além disso, temos os famosos 5 Vs do Big Data, que representam o volume, velocidade, variedade, valor e veracidade dos dados. Um cientista de dados precisa ter conhecimento em Estatística, Machine Learning, Linguagens de Programação, Banco de Dados, Tecnologias de Informação. Como o Big Data exige técnicas estatísticas avançadas, os profissionais costumam vir de cursos como matemática industrial, engenharia, estatística e ciência da computação, com pós-graduação voltada para o estudo dos dados. As habilidades desenvolvidas por um profissional da área são: comunicação, gestão de dados estruturados, gestão de projetos, data mining, desenvolvimento de produtos, modelagem estatística e desenvolvimento de negócios, enquanto as principais competências são: curiosidade, atualização constante, comunicação e pensamento crítico.

Palavras-chave: Big Data. Data Science. Banco de dados. Mineração de dados.