

Ronaldo Goldschmidt

Possui graduação em Matemática pela Universidade Federal Fluminense, mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia e doutorado em Engenharia Elétrica - Métodos de Apoio à Decisão pela Pontifícia Universidade Católica do Rio de Janeiro. Atualmente é professor adjunto do Instituto Militar de Engenharia. Tem experiência na área de Ciência da Computação, com ênfase em Inteligência Artificial e Lógica, atuando principalmente nos seguintes temas: sistemas de apoio à decisão, mineração de dados e textos, linguagens formais e banco de dados.

Emmanuel Passos

É Professor Titular e Doutor em Computação pela COPPE/UFRJ, com experiência em ensino universitário desde 1968. Atualmente é professor colaborador do Departamento de Engenharia Elétrica na PUC-Rio e consultor de empresas, tendo desenvolvido diversas aplicações de Data Mining no mercado.

Eduardo Bezerra

Possui graduação em Ciência da Computação pela UFRJ, mestrado e doutorado em Engenharia de Sistemas e Computação pela COPPE/UFRJ. Desde 2005, é professor da Escola de Informática e Computação do CEFET/RJ. Seus interesses de pesquisa incluem Engenharia de Software, Linguagens de Programação e Mineração de Dados.

Consulte nosso catálogo completo e últimos lançamentos em:
www.elsevier.com.br

Os constantes avanços na área da Tecnologia da Informação têm viabilizado o armazenamento de grandes, múltiplas, heterogêneas e distribuídas bases de dados de natureza administrativa, científica, comercial, educacional, governamental e social em todo o mundo. Estima-se que, em 2020, a humanidade terá cerca de 44 zettabytes de informações digitais disponíveis. Na era globalizada em que vivemos onde é cada vez maior a competitividade entre as empresas, informação e conhecimento são elementos fundamentais para se obter diferenciais mercadológicos frente à concorrência. Então, como tirar proveito do conhecimento subjacente a todo patrimônio digital que vem sendo acumulado? A análise desses dados pelo homem é inviável sem o auxílio de ferramentas computacionais apropriadas.

Diante deste cenário, encontra-se uma nova área popularmente conhecida como *Big Data*, para referenciar (e lidar com) esse massivo e acelerado crescimento de volumes de dados ao redor do mundo. *Data Mining*, foco deste livro, é o subconjunto do universo *Big Data* que se concentra no desenvolvimento e na aplicação de técnicas que permitam analisar e obter conhecimentos novos e úteis a partir de grandes bases de dados.

Diversas técnicas de *Data Mining* vêm sendo utilizadas com sucesso no mundo inteiro e estão presentes em várias instituições (nacionais e internacionais) tais como: Nasa, Walmart, FedEx, UPS, Amazon, IBM, Caixa Econômica Federal, Banco do Brasil, Vale do Rio Doce, Petrobras, Dataprev, Serpro, EMC, entre muitas outras. Detecção de fraudes em arrecadações, tendências de consumo e de opiniões de clientes, acompanhamento personalizado de estudantes, compreensão da evolução de redes sociais virtuais, previsão de produção na indústria e de riscos no mercado financeiro são alguns exemplos dentre as inúmeras aplicações de Data Mining na atualidade.

Assim sendo, este livro se destina a todas as pessoas relacionadas à área de TI (analistas, programadores, gerentes, professores, estudantes) que desejem atuar em *Data Mining*, analisando dados e tirando proveito do conhecimento existente em conjuntos de dados nos mais variados contextos de aplicação.



ELSEVIER

Ronaldo Goldschmidt | Emmanuel Passos | Eduardo Bezerra

Data Mining

CAMPUS



Ronaldo Goldschmidt
Emmanuel Passos
Eduardo Bezerra

Data Mining

2ª
edição

Conceitos, técnicas,
algoritmos, orientações
e aplicações

Material
na WEB

www.elsevier.com.br

Recursos na web:

Como informação suplementar à contida neste livro, disponibilizamos um site na própria editora Elsevier/Campus. O leitor pode acessar a página da editora (www.elsevier.com.br). Nesse endereço, o leitor pode obter informações e material complementar sobre o tema. O leitor pode também utilizar esse site para entrar em contato com os autores, com o objetivo de trocar ideias sobre o livro. Entre os recursos que podem ser encontrados no site, estão os seguintes:

- Soluções de alguns dos exercícios propostos no livro;
- Apresentações baseadas no conteúdo dos assuntos abordados no livro. Esse material é útil para o professor ou instrutor que deseja adotar o livro em seus cursos.;
- Implementações de alguns dos métodos e/ou técnicas apresentados no livro;
- Um texto introdutório sobre Redes Neurais Artificiais, Lógica Nebulosa, Algoritmos Genéticos e Data Warehouse. Recomenda-se que leitores não familiarizados com os temas, leiam esse texto antes do Capítulo 5 do livro;
- Outras fontes de informação. O material disponível no site da editora contém também endereços para outras fontes de informação sobre Data Mining. Seguindo a natureza dinâmica da internet, o conteúdo do site será modificado de tempos em tempos.

Data ● Mining