

Fundamentos de Ciência de Dados e *Business Analytics*

Victor F. Cavalcante

IBM Research – Brasil

MC959A/MO810B

Objetivos

- Formar conceito sobre o que vem a ser Analytics.
- Visão geral de tópicos em Analytics/BAO (Business Analytics & Optimization).
- Overview de áreas envolvidas e ilustrações cobrindo alguns tópicos de cada uma das disciplinas.
- Prover visão de mercado em Analytics & Data Science.
- Ilustrar situações de pesquisa aplicada (e multidisciplinar) relacionada.
- Desenvolvimento de trabalho em time (grupo com dinâmica de indústria):
 - Brainstorming sessions/discussões;
 - Atribuição de tarefas e papéis;
 - Apresentação de resultados;
 - Competitividade;
 - Julgamento de mérito e possibilidade de divulgação/publicação de resultados.

Tópicos & Organização

- Analytics na Indústria:
 - Cenários/casos ilustrativos
- Aspectos/Tópicos em:
 - Learning;
 - Big Data;
 - (Social & Data) Mining;
 - Algoritmos de Otimização, e
 - Computação Cognitiva.
- Dinâmica/formato:
 - Aulas expositivas (aprox. 70% do curso);
 - Apresentações de videos, incluindo depoimentos de expoentes na área;
 - Leituras de textos de revistas (ORMS e Analytics) e livros (discussões);
 - Apresentação de casos/cenários da indústria;
 - Descrições de aplicações e projetos;
 - Aprendizado “em ciclos”.
 - Contribuições externas (workshops e palestras: tirem proveito!!!).

Bibliografia

- Y. Abu-Mostafa et al., *Learning From Data – A short Course*, AMLBook (2012)
- J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Elsevier (2006).
- M. Gendreau and J. Potvin, **Handbook of Metaheuristics**, 2nd Edition, Int. Series in Operations Research & Management Science, Vol. 146, Springer (2010).

Complementar:

- Conteúdos de reportagens das revistas ***Analytics*** e ***ORMS Today*** do INFORMS (Institute for Operations Research and the Management Sciences).
- J. Leskovec, A. Rajaraman and J.D. Ullman, *Mining of Massive Datasets*, Cambridge University Press (2012).
- F. Glover and G. A. Kochenberger, *Handbook of Metaheuristics*, INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE, KLUWER Academic Publishers (2003).
- Bing Liu, *Web Data Mining – Exploring Hyperlinks, Contents and Usage Data*, 2nd Edition, Springer (2011)
- D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, Cambridge University Press (2010).
- C.C. Aggarwal (Editor), *Social Network Data Analytics*, Springer New York (2011).
- M. E. J. Newman, *Networks: An Introduction*, Oxford University Press (2010).

Avaliação

- Atividades:
 - Pesos diferentes
 - Projeto (trabalho em grupo):
 - Apresentação de seminário temático (N_1)
 - Implementação (algoritmo) (N_2)
 - Lista de exercícios (N_3)
 - Estudos e discussões sobre textos (N_4)
 - Modelagem com SPSS* (N_5)
- Nota Final: $2 \times N_1 + 4 \times N_2 + N_3 + N_4 + 2 \times N_5$
 - * O uso do SPSS está condicionado à liberação de licença para o IC/UNICAMP, para uso na disciplina (solicitação em andamento). Caso não se confirme, a atividade será suprimida e, ambas, N_3 e N_4 , terão peso 2 na nota final.
- Prazos:
 - A serem definidos até 17/09/2014 (após alterações de matrícula).
 - Atividades entregues após o prazo estipulado sofrerão uma penalidade de 50% da nota. Entregas com mais de 1 semana (7 dias corridos) de atraso não serão aceitas; ou seja, será atribuída nota ZERO para uma atividade com mais de 7 dias de atraso.

Horário de Atendimento/Ausências

- Às 4as-feiras, após o término das aulas.
- Via e-mail: victorf.cavalcante@gmail.com
- 2 viagens a trabalho (conferências) já programadas:
 - *08/09/2014 e 10/09/2014*
 - *10/11/2014 e 12/11/2014*

Nota: Atividades em grupo serão programadas para cobrir parte do período ausente.