



COMPONENTE CURRICULAR

CÓDIGO	NOME	PRÉ-REQUISITO
MATF52	Introdução à Mineração de Dados	MAT229 – Análise de Regressão

CARGA HORÁRIA				MÓDULO			SEMESTRE VIGENTE
T	P	E	TOTAL	T	P	E	
34	34	00	68	40	40	00	A partir de 2018.2

EMENTA

Introdução à mineração de dados. Análise estatística de grandes bancos de dados. O processo de descoberta de conhecimento em bancos de dados. Principais funcionalidades e técnicas de mineração de dados. Classificação de dados. Métodos de classificação supervisionada. Combinação de classificadores. Segmentação. Estudos de caso.

OBJETIVOS

CONTEÚDO PROGRAMÁTICO

1. Introdução à mineração de dados.
 - 1.1. O que é mineração de dados?
 - 1.2. A interdisciplinaridade da técnica.
 - 1.3. Problemas típicos e potenciais aplicações de mineração de dados.
2. Mineração de dados e análise estatística.
 - 2.1. Entendendo as diferenças.
 - 2.2. Questionando estratégias.
3. Importância dos bancos de dados.
 - 3.1. Características em um banco de dados.
 - 3.2. As idéias de armazém de dados (*Data Warehouse*) e processamento analítico em tempo real (OLAP).
4. Visão geral sobre as fases do processo de descoberta de conhecimento em bancos de dados (*Knowledge Discovery in Databases*).
5. Classificação supervisionada.
 - 5.1. Métodos.
 - 5.1.1. Árvores de decisão.
 - 5.1.2. Regressão logística.
 - 5.1.3. Naive Bayes.
 - 5.1.4. K vizinhos mais próximos.
 - 5.2. Avaliação dos modelos de classificação supervisionada.
 - 5.2.1. Medidas de capacidade preditiva.
 - 5.2.1.1. Sensibilidade.
 - 5.2.1.2. Especificidade.

- 5.2.1.3. Valores de predição positivo e negativo.
- 5.2.1.4. Acurácia.
- 5.2.2. Curva ROC.
- 6. Combinação de classificadores.
 - 6.1. *Bagging* de modelos.
 - 6.2. Métodos de combinação.
 - 6.2.1. Combinação via voto.
 - 6.2.2. Combinação via regressão logística.
- 7. Procedimentos de segmentação de base.
 - 7.1. Segmentação por benefícios.
 - 7.2. Segmentação demográfica.
 - 7.3. Segmentação por nível de utilização de um produto.
 - 7.4. Segmentação por estilo de vida.
 - 7.5. Segmentação com base no valor do cliente.

BIBLIOGRAFIA

Bibliografia Básica

- BRAMER, M. Principles of Data Mining. London: Springer-Verlag, 2ª edição, 2013, 440 p. (ebook)
- JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. An Introduction to Statistical Learning: with Applications in R. New York: Springer-Verlag, 2013, 426 p. (ebook)
- JOHNSON, R. A.; WICHERN, D. W. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall, 4ª edição, 1998, 816 p.
- KUHN, M.; JOHNSON, K. Applied Predictive Modeling. New York: Springer-Verlag, 2013, 600 p. (ebook)
- SANTOS, L. A. N. Contribuição da mineração de dados e da otimização heurística para a interpretação dos dados da produção científica brasileira. 2011. 114 f. Dissertação (Mestrado em Ciência da Informação) – Instituto de Ciência da Informação, Universidade Federal da Bahia, Salvador. 2011.

Bibliografia Complementar

- CUZZOCREA, A.; DAYAL, U. (eds.). Data Warehousing and Knowledge Discovery. 14th International Conference, DaWaK 2012, Vienna, Austria, September 3-6, 2012, Proceedings. Berlin/Heidelberg: Springer-Verlag, 454 p. (ebook)
- DOUGHERTY, G. Pattern Recognition and Classification: An Introduction. New York: Springer-Verlag, 2013, 196 p. (ebook)
- GABER, M.M. Journeys to Data Mining: Experiences from 15 Renowned Researchers. Berlin/Heidelberg: Springer-Verlag, 2012, 244 p. (ebook)
- GIUSTI, A.; RITTER, G.; VICHI, M. (eds.). Classification and Data Mining. *Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin/Heidelberg: Springer-Verlag, 2013, 286 p. (ebook)
- KLEINBAUM, D. G. Logistic regression: a self-learning text. *Statistics in the health sciences*. New York: Springer-Verlag, 1994, 282 p.
- MICHIE, D.; SPIEGELHALTER, D.J.; TAYLOR, C.C. Machine Learning, Neural and Statistical Classification. New York: Ellis Horwood, 1994, 289 p.

Teses, dissertações, artigos e relatórios técnicos que apresentem metodologias de Mineração de Dados.

Assinatura e Carimbo do Chefe do Departamento
Programa aprovado em reunião plenária do dia 01/09/2017

Assinatura e Carimbo do Coordenador do Curso
Programa aprovado em reunião plenária do dia ___/___/___