

Técnicas Avançadas de Captura e Tratamento de Dados

Apresentação

Moacir Antonelli Ponti
ICMC/USP

moacir@icmc.usp.br

Luis Gustavo Nonato
ICMC/USP

gnonato@icmc.usp.br



Tópicos

1. Dados estruturados e Não Estruturados e problemas típicos em bases de dados reais
2. Tratamento da qualidade dos dados, valores faltantes, ruídos, outliers e redundância
3. Transformação dos dados: técnicas de discretização, normalização e transformação
4. Mudança de representação de dados: redução de dimensionalidade, transformação e seleção de atributos

Tópicos

5. Dados não estruturados I - capturando texto
6. Dados não estruturados II - estruturando texto
7. Dados não estruturados III - representações de sinais e imagens
8. Web Crawling via python Request

Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd edition O'Reilly Media, 2017

Mark Lutz, Learning Python, 5th Edition, O'Reilly Media, 2013

Eric Matthew, Python Crash Course: a hands-on, project-based introduction to programming, no starch press, 2013

Katti Faceli, Inteligência Artificial. Uma Abordagem de Aprendizado de Máquina, LTC 2012

Al Sweigart, Automate the Boring Stuff with Python: Practical Programming for Total Beginners, Starch Press 2015