



Serviço Público Federal
Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense
Pró-reitoria de Ensino

DISCIPLINA: Introdução à Ciência de Dados	
Vigência: a partir de 2021/1	Período letivo: ELETIVA
Carga horária total: 45 h	Código: SUP.2280
Ementa: Introdução ao estudo a ciências de dados; questões éticas e de privacidade na coleta dos dados. Estudo de técnicas para coleta e pré-processamento de dados. Fundamentação sobre análise exploratória de dados; introdução a mineração de dados e desenvolvimento prático de casos de estudo.	

Conteúdos:

UNIDADE I - Introdução a ciência de dados

- 1.1. O ecossistema de ciência dos dados
- 1.2. Informação VS conhecimento
- 1.3. Ética e privacidade no contexto de ciência dos dados

UNIDADE II - Preparação dos dados para análise

- 2.1. Coleta de dados (data scraping)
- 2.2. Técnicas de Pré-processamento dos dados

UNIDADE III - Análise exploratória de dados

- 3.1. Conceitos de análise exploratória
- 3.2. Estatísticas descritivas aplicada a ciência dos dados
- 3.3. Visualização de dados

UNIDADE IV - Introdução a mineração de dados

- 4.1. Aprendizado estatístico
- 4.2. Classificação
- 4.3. Agrupamento
- 4.4. Mineração de padrões frequentes
- 4.5. Medidas de interesse/qualidade

UNIDADE V - Desenvolvimento de estudo de caso

- 5.1. Especificação de estudo de caso
- 5.2. Avaliação do estudo de caso

Bibliografia básica

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; CARVALHO, André.
Inteligência Artificial: Uma abordagem de aprendizado de máquina. São Paulo, SP: Grupo Gen-LTC, 2011.



Serviço Público Federal
Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense
Pró-reitoria de Ensino

GRUS, Joel. **Data Science do zero: Primeiras regras com o Python**. Rio de Janeiro, RJ: Alta Books, 2019.

BARBETTA, Pedro Alberto; REIS, Marcelo Menezes; BORNIA, Antonio Cezar. **Estatística: para cursos de engenharia e informática**. 3. ed. São Paulo, SP: Atlas, 2010. 410 p. ISBN 9788522459940.

Bibliografia complementar

RUSSELL, Stuart; NORVIG, Peter (Aut.). **Inteligência artificial**. Rio de Janeiro, RJ: Elsevier, 2004. 1021 p. ISBN 9788535211771.

RUSSELL, Matthew A. **Mineração de dados da web social**. São Paulo, SP: Novatec, 2011. 357 p. ISBN 9788575222454.

RAMAKRISHNAN, Raghu; GEHRKE, Johannes (Aut.). **Sistemas de gerenciamento de banco de dados**. 3. ed. São Paulo, SP: McGraw-Hill, 2008. xxvii, 884 p. ISBN 8577260270.

BARBETTA, Pedro Alberto; REIS, Marcelo Menezes; BORNIA, Antonio Cezar. **Estatística: para cursos de engenharia e informática**. 3. ed. São Paulo, SP: Atlas, 2010. 410 p. ISBN 9788522459940.

HAYKIN, Simon S. **Redes neurais/princípios e práticas**. 2. ed. Porto Alegre, RS: Bookman, 2001. xxv, 900 p. ISBN 9788573077186.