



PROGRAMAS E BIBLIOGRAFIAS

DISCIPLINA AP528	NOME DA DISCIPLINA PORTUGUÊS – CIÊNCIA DOS DADOS I: INTRODUÇÃO INGLÊS – DATA SCIENCE I: INTRODUCTION ESPAÑHOL - LA CIENCIA DE DATOS I: INTRODUCCIÓN
----------------------------	---

Horas Semanais							
Aulas Teóricas	15	Aulas Práticas	30	Aulas de Estudo Dirigido	0	Horas em Sala de Aula	45
Nº semanas	Carga horária total	Créditos	Frequência		Aprovação		
15	45	3	75 %		Nota e frequência		

Ementa:

Redução de dimensionalidade dos dados: Amostragem e seleção de atributos. Técnicas para balanceamento dos dados. Preenchimento de dados ("Imputação") e remoção de "outliers". Seminários de tópicos avançados. Projetos com aplicações na agricultura.

Objetivos:

Explorar as técnicas existentes para preparar os dados visando a fase de modelagem, aprendendo a utilizá-las de forma adequada conforme o problema apresentado.

Programa:

Parte I – Qualidade dos Dados:

- Valores Faltantes e Técnicas de Imputação
- 'Outliers', Ruidos e Integração (Inconsistência/Conflitos/Redundância)

Parte II – Transformação e Sintetização dos Dados

- Normalização (min-max, z-score), 'smoothing', Discretização, Histogramas, etc.

Parte III – Redução da Dimensionalidade

- Amostragem
- Seleção de Atributos

Bibliografia:

Referências básicas:

- MIAO, J.; NIU, L. (2016) A Survey on Feature Selection. Procedia Computer Science 91, 919-926.
- CHANDRASHEKAR, G.; SAHIN, F. (2014) A survey on feature selection methods. Computers and Electrical Engineering 40 (2014) 16–28.
- LIU, H.; MOTODA, H. (Eds.) (2008) Computational Methods of Feature Selection. Chapman & Hall/CRC.
- DASU, T.; JOHNSON, T. (2003) Exploratory Data Mining and Data Cleaning. Wiley.
- RODERICK, J.L.; RUBIN, D.B. (2002) Statistical analysis with missing data. Wiley.
- PYLE, D. (1999). Data Preparation for Data Mining (1st Ed.). San Francisco, California, USA: Morgan Kaufmann.

Referências Complementares:

- TAN, P.-N., STEINBACH, M., & KUMAR, V. (2009). Introdução à Mineração de Dados (1st ed.). Editora Ciência Moderna
- MUCHERINO, A., PAPAJOJGI, P. J., & PARDALOS, P. M. (2009). Data mining in agriculture (1st ed.). Springer Verlag

Critérios de Avaliação:

- Seminários, Resenhas e Análises Críticas de artigos
- Projeto

Observações:

ASSINATURAS:

Prof. Dr. Luiz Henrique A. Rodrigues
Matrícula 130389
FEAGRI/UNICAMP

Instruções:

Horas em sala - Ao total de horas de aulas teóricas ministradas (T) deverá obrigatoriamente ser associada carga horária em sala de aula.

EMISSÃO: 10 / 05 / 2019
PÁGINA: 1 de 1

Rubrica: