



**CURSO:** Programa de Pós-graduação em Engenharia de Bioprocessos e Biotecnologia

**Nível:** Mestrado

### INFORMAÇÕES BÁSICAS

<b>Unidade curricular</b> Práticas em Mineração de Dados para não programadores				<b>Unidade</b> PPGEnBio
<b>Carga Horária</b>				<b>Código</b> Preenchido pela secretaria
<b>Teórica</b> 15	<b>Prática</b> 30	<b>Total</b> 45	<b>Créditos</b> 3	
<b>Tipo</b> Optativa	<b>Habilitação / Modalidade</b> Mestre / Doutor		<b>Pré-requisito</b> -	

### EMENTA

Introdução. Entrada: conceitos, instâncias, atributos. Saída: representação do conhecimento. Algoritmos: os métodos básicos. Credibilidade: Avaliando o que foi aprendido. Preparação e análise exploratória de dados.

### OBJETIVOS

Apresentar os principais conceitos sobre Mineração de dados e capacitar os alunos na utilização do framework Weka para aplicação de Mineração de Dados nos mais variados problemas.

### CONTEÚDO PROGRAMÁTICO

- **1. Introdução:**
  - 1.1 Mineração de Dados e Aprendizado de Máquina;
  - 1.2 Exemplos Simples: O Problema do Clima e Outros;
  - 1.4 O Processo de Mineração de Dados;
  - 1.5 Aprendizado de Máquina e Estatística;
  - 1.6 Generalização como Busca.
- **2. Entrada: conceitos, instâncias, atributos**
  - 2.1 O que é um Conceito?
  - 2.2 O que há em um Exemplo?
  - 2.3 O que há em um Atributo?
  - 2.4 Preparando a Entrada.
- **3. Saída: representação do conhecimento**
  - 3.1 Tabelas;
  - 3.2 Modelos Lineares;
  - 3.3 Árvores;
  - 3.4 Regras;
  - 3.5 Representação Baseada em Instâncias;
  - 3.6 Agrupamentos.
- **4. Algoritmos: os métodos básicos**
  - 4.1 Inferindo Regras Rudimentares;

<ul style="list-style-type: none"> <li>○ 4.2 Modelagem Probabilística Simples;</li> <li>○ 4.3 Dividir para Conquistar: Construindo Árvores de Decisão;</li> <li>○ 4.4 Algoritmos de Cobertura: Construindo Regras;</li> <li>○ 4.5 Mineração de Regras de Associação;</li> <li>○ 4.6 Modelos Lineares;</li> <li>○ 4.7 Aprendizado Baseado em Instâncias;</li> <li>○ 4.8 Agrupamento.</li> <li>● <b>5. Credibilidade: Avaliando o que foi aprendido</b> <ul style="list-style-type: none"> <li>○ 5.1 Treinamento e Teste;</li> <li>○ 5.2 Prevendo o Desempenho;</li> <li>○ 5.3 Validação Cruzada;</li> <li>○ 5.4 Outras Estimativas;</li> <li>○ 5.5 Seleção de Hiperparâmetros;</li> <li>○ 5.6 Comparando Esquemas de Mineração de Dados.</li> </ul> </li> <li>● <b>6. Preparação e análise exploratória de dados</b> <ul style="list-style-type: none"> <li>○ 6.1 Seleção de Atributos;</li> <li>○ 6.2 Discretizando Atributos Numéricos;</li> <li>○ 6.3 Projeções;</li> <li>○ 6.4 Amostragem;</li> <li>○ 6.5 Limpeza;</li> <li>○ 6.6 Transformando Múltiplas Classes em Binárias;</li> <li>○ 6.7 Calibrando Probabilidades de Classe;</li> <li>○ 6.8 Análise Exploratória de Dados.</li> </ul> </li> </ul>
<b>CRITÉRIOS DE AVALIAÇÃO</b>
<ol style="list-style-type: none"> <li>1. Apresentação de artigo – 25 pontos;</li> <li>2. Realização de 3 trabalhos práticos em sala de aula, utilizando o framework Weka – 25 pontos cada trabalho.</li> </ol> <p><b>- Será considerado aprovado o aluno que obtiver o mínimo de 60% da pontuação distribuída.</b></p>
<b>BIBLIOGRAFIA BÁSICA</b>
<ol style="list-style-type: none"> <li>1. MARIANO, Diego César Batista; MARQUES, Leonardo Torres; SILVA, Marcel Santos. Data mining. 1. ed. Porto Alegre: SAGAH, 2021. 1 recurso online (0 p.). ISBN 9786556900292</li> <li>2. ZAKI, Mohammed J.; MEIRA, Wagner. Data mining and machine learning: fundamental concepts and algorithms. 2nd ed. Cambridge, UK; New York: Cambridge University Press, 2020. 766 p. ISBN 9781108473989 (enc.).</li> <li>3. GOLDSCHMIDT, Ronaldo. Data mining. 2. ed. Rio de Janeiro: LTC, 2015. 1 recurso online (344 p.). ISBN 9788595156395.</li> </ol>
<b>BIBLIOGRAFIA COMPLEMENTAR</b>
<ol style="list-style-type: none"> <li>1. TAN, Pang-ning; STEINBACH, Michael; KUMAR, Vipin. Introdução ao data mining: mineração de dados. Rio de Janeiro: Ciência Moderna, 2009 900 p.</li> <li>2. BARBIERI, Carlos. Governança de dados. 1. ed. Rio de Janeiro: Alta Books, 2020. 1 recurso online (0 p.). ISBN 9788550815435.</li> <li>3. CARVALHO, André C. P. L. F. de; MENEZES, Angelo Garangau; BONIDIA, Robson Parmezan. Ciência de dados. 1. ed. Rio de Janeiro: LTC, 2024. 1 recurso online (0 p.). ISBN 9788521638766.</li> </ol>



UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI – UFSJ  
INSTITUÍDA PELA LEI Nº.10.425 DE 19/04/2002 – D.O.U. DE 22/04/2002



**Programa de Pós-graduação em Eng. de Bioprocessos e Biotecnologia**  
**Campus avançado Alto Paraopeba**

Rodovia MG 443, Km 7, Ouro Branco/Minas Gerais/Brasil

4. GRUS, Joel. Data science do zero. 2. ed. Rio de Janeiro: Alta Books, 2021. 1 recurso online (0 p.). ISBN 9788550816463.
5. BRUCE, Peter; BRUCE, Andrew. Estatística prática para cientistas de dados: 50 conceitos essenciais. 1. ed. Rio de Janeiro: Alta Books, 2019. 1 recurso online (0 p.). ISBN 9788550813004.