



UNIVERSIDADE FEDERAL DE OURO PRETO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PLANO DE ENSINO



Nome do Componente Curricular em português: Mineração de Dados Nome do Componente Curricular em inglês: Data Mining		Código: PCC142
Nome e sigla do departamento: Departamento de Computação (DECOM)		Unidade acadêmica: ICEB
Nome do docente: Anderson Almeida Ferreira		
Carga horária semestral 60 horas	Carga horária semanal teórica 4 horas/aula	Carga horária semanal prática 0 horas/aula
Data de aprovação: / 03 / 2025		
Ementa: Introdução e aplicações de mineração de dados. Processo de descoberta de conhecimento em bases de dados. Pré-processamento de dados. Técnicas de mineração de dados para as tarefas de extração de regras de associação e padrões de sequência, agrupamento (clustering), classificação e regressão. Uso de ferramentas de mineração de dados.		
Conteúdo programático: 1. Introdução e aplicações de mineração de dados. 1.1. Técnicas e tarefas de mineração de dados. 1.2. Exemplos de aplicações. 2. Processo de descoberta de conhecimento em bases de dados. 2.1. Descrição das etapas do processo de descoberta de conhecimento em bases de dados. 3. Pré-processamento de dados. 3.1. Limpeza de dados. 3.2. Integração de dados. 3.3. Transformação de dados. 3.4. Redução de dados. 3.5. Discretização de dados. 4. Extração de regras de associação e padrões de sequência. 4.1. Conceitos iniciais de regras de associação e padrões de sequência. 4.2. Tipos de regras de associação. 4.3. Modelo suporte/confiança. 4.3. Algoritmos Apriori e Partition. 4.4. Outras medidas de interesse. 4.5. Algoritmo AprioriAll. 5. Classificação e regressão. 5.1. Processo de classificação. 5.2. Classificação por indução de árvores de decisão. 5.3. Classificadores bayesianos.		

- 5.4. Classificador k-NN.
- 5.5. Avaliação de classificadores.
- 5.6. Regressão linear simples e múltipla.
- 6. Agrupamento (clustering).
 - 6.1. Conceitos iniciais.
 - 6.2. Funções de distância.
 - 6.3. Preparação dos dados.
 - 6.4. Categorias de métodos de agrupamento.
 - 6.5. Algoritmo k-means.
- 7. Uso de ferramentas de mineração de dados.
 - 7.1. Exercícios envolvendo o uso de ferramentas de mineração

Objetivos:

Apresentar aos alunos diversos aspectos teóricos e práticos de mineração de dados.

Metodologia:

Aulas expositivas sobre o conteúdo programático.

Exercícios: Realização de exercícios para fixação de assuntos estudados e utilização de ferramentas de mineração de dados.

Trabalho prático: Implementação e análise de problemas usando técnicas de Mineração de Dados. Será avaliado por meio de apresentação e avaliação de relatório/artigo.

Leituras recomendadas: leitura de textos técnicos com a finalidade de proporcionar ao discente a oportunidade de consulta e desenvolvimento de sua capacidade de análise, síntese e crítica de uma bibliografia específica.

Seminários: Apresentação de artigos relacionados a temas da disciplina.

Recursos:

Moodle, Google Meet, Kaggle Datasets, Jupyter Notebook, Google Colaboratory

Atividades avaliativas:

Exercícios valendo 10 pontos com peso 1,5.

Seminários valendo 10 pontos com peso 3,5.

Trabalho Prático valendo 10 pontos com peso 5. O trabalho prático será dividido em 3 fases: 1) Pré-processamento e transformação (25%); 2) Mineração de padrões frequentes (25%); 3) Mineração para atingir os objetivos do desafio, avaliação, análise de resultados iniciais, melhoria dos resultados da mineração inicial e artigo com resultados finais (50%).

Cronograma:

	Semana	Tópicos
1	17/03 e 19/03	Apresentação da disciplina. Introdução a mineração de dados: Aplicações de Mineração de Dados; Processo de descoberta de conhecimento em bases de dados
2	24/03 e 26/03	Pré-processamento de dados: Limpeza de dados, Transformação de Dados, Redução de Dados e Integração de Dados Seleção de atributos

3	31/03 e 02/04	Entrega da proposta do trabalho prático a ser desenvolvido Seminários
4	07/04 e 09/04	Regras de associação
5	14/04	Mineração de padrões de sequência
6	05/05 e 07/05	Seminários
7	12/05 e 14/05	Entrega da primeira fase do trabalho prático Apresentação da primeira fase do trabalho prático
8	19/05 e 21/05	Classificação – Introdução; Árvores de decisão; Bayesianos; KNN
9	26/05 e 28/05	Classificação – SVM; Backpropagation; Avaliação
10	02/06 e 04/06	Entrega da segunda fase do trabalho prático Apresentação da segunda fase do trabalho prático
11	09/06 e 11/06	Seminários
12	16/06 e 18/06	Agrupamento/Segmentação
13	23/06 e 25/06	Avaliação de Agrupamento/Segmentação
14	30/06 e 02/07	Seminários
15	07/07 e 09/07	Entrega da terceira fase do trabalho prático Apresentação da terceira fase do trabalho prático
16	14/07	Apresentação da terceira fase do trabalho prático

Bibliografia básica:

Mohammed J. Zaki, Wagner Meira, Jr., **Data Mining and Machine Learning: Fundamental Concepts and Algorithms**, 2nd Edition, Cambridge University Press, March 2020. ISBN: 978-1108473989. Disponível em: https://dataminingbook.info/book_html/

Silva, L. A., Peres, S. M, Boscarioli, C.. **Introdução à Mineração de Dados - Com Aplicações em R**. Editora Elsevier: Grupo GEN, 2016. 9788595155473. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595155473/> .

Castro, L. N.; Ferrari, D. G.. **Introdução à mineração de dados : conceitos básicos, algoritmos e aplicações**. 1ª. Edição. São Paulo: Saraiva, 2016. ISBN 978-85-472-0100-5. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/>

Han, J., Kamber, M., **Data Mining: Concepts and Techniques**, Morgan Kaufmann Publishers, 2a. edição, 2006. ISBN: 1558609016

Bibliografia complementar:

Goldschmidt, R.; Passos, E. L. **Data Mining: um guia prático : conceitos, técnicas, ferramentas, orientação e aplicações** . Rio de Janeiro: Elsevier, Campus: Grupo GEN, 2015. 9788595156395. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156395/> .

Faceli, K., Lorena, A.C., Gama, J., Carvalho, A.C.P.L.F. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina**. LTC: Grupo GEN, 2011. 978-85-216-2146-1. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-216-2146-1/> .

Norvig, P. **Inteligência Artificial**. Campus: Grupo GEN, 2013. 9788595156104. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156104/> .

Sharda, R., Delen, D., & Turban, E. **Business Intelligence e Análise de Dados para Gestão do Negócio**. Bookman, 2019. 9788582605202. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788582605202/>.

VanderPlas, J. Python Data Science Handbook. O'Reilly Media, 2016. 9781491912058. Disponível em: <https://colab.research.google.com/github/jakevdp/PythonDataScienceHandbook/blob/master/notebooks/Index.ipynb#scrollTo=RhZ1h21YsWnM>.

Witten, I.H., Frank, E., **Data Mining: Practical Machine Learning Tools and Techniques**. Morgan Kaufmann, 2a. edição, 2005. ISBN: 0120884070.

Linoff, G.S., Berry, M.J.A., **Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management**, Wiley, 3a edição, 2011. ISBN: 0470650931.

LIU, Bing. **Web data mining: exploring hyperlinks, contents, and usage data**. Berlin: New York: Springer c2007. xix, 532 p. (Data-centric systems and applications). ISBN 9783540378815.