



UNIVERSIDADE FEDERAL DE OURO PRETO  
PRÓ-REITORIA DE GRADUAÇÃO  
PLANO DE ENSINO



Nome do Componente Curricular em português: <b>Introdução à Mineração de Dados</b>		Código:
Nome do Componente Curricular em inglês: Introduction to Data Mining		EST125
Nome e sigla do departamento: Departamento de Estatística DEEST		Unidade acadêmica: ICEB
Nome do docente: Tiago Martins Pereira		
Carga horária semestral 60 horas	Carga horária semanal teórica 04 horas/aula	Carga horária semanal prática 00 horas/aula
Data de aprovação na assembleia departamental: / / 2024		
Ementa: Definição. Relação com a descoberta de conhecimento, estatística e inteligência computacional. Obtenção, normalização e limpeza de dados. Seleção, transformação e mineração. Classificação e regressão. Modelos e algoritmos para mineração de dados. Extração de informações e síntese. Classificação supervisionada, não-supervisionada e algoritmos híbridos. Árvores de decisão e sua criação. Algoritmos e técnicas de regressão. Interação entre técnicas de mineração de dados e outras técnicas computacionais.		
Conteúdo programático		
<b>Unidade 1:</b> Definição de Mineração de Dados. Relação do processo de mineração de dados com descoberta de conhecimento, estatística, inteligência computacional. Fases de um processo de descoberta de conhecimento: obtenção e normalização de dados, limpeza de dados, seleção e transformação, mineração, avaliação do conhecimento.		
<b>Unidade 2:</b> Análise preditiva - Tarefas de mineração de dados: classificação e regressão.		
<b>Unidade 3:</b> Análise de agrupamentos: A tarefa de agrupamento. Avaliação de modelos para análise de agrupamento.		
<b>Unidade 4:</b> Regras de associação: Processo de mineração de regras de Associação. Algoritmos para mineração de regras de associação.		
<b>Unidade 5:</b> Detecção de anomalias: Introdução. Processo de detecção de anomalias. Métodos estatísticos. Métodos algorítmicos.		
Objetivos:		
Fornecer ao aluno a teoria e modelo de cada uma das principais técnicas de Mineração de Dados no enfoque estatístico, bem como mostrar as aplicações práticas nos diversos ramos da ciência destas técnicas. O aluno deverá ser capaz de utilizar os programas de computador relacionados com as técnicas apresentadas.		

## Metodologia:

Aulas expositivas sobre o conteúdo programático e textos para estudo.

## Atividades avaliativas:

O atendimento aos objetivos da disciplina será avaliado através de atividades avaliativas online. As atividades consistirão em trabalhos práticos aplicados à resolução de problemas contextualizados aos diferentes tópicos do conteúdo programático, entregues em formato digital, através da plataforma Moodle.

- Os pontos serão distribuídos da seguinte forma:

- 5 (cinco) trabalhos práticos (distribuídos ao longo do curso) - VALOR: 100,0 pontos

Será aprovado o aluno que atingir média final igual ou superior a 60,0 pontos. Para não ser reprovado(a) por infrequência e ter possibilidades de realizar o exame especial (caso necessário), o(a) aluno(a) deverá ter pelo menos 75% de frequência nas atividades da disciplina.

Alunos com média final inferior a 60,0 pontos poderão ser aprovados através do Exame Final versando sobre toda a matéria tratada na disciplina.

**Obs.:** as atividades inerentes à disciplina “Introdução à mineração de dados” são de extrema importância para o desenvolvimento ideal do tema.

Todas as atividades propostas nesse curso valem pontuação. Assim sendo, todas as atividades serão pontuadas, somando 100,0 pontos distribuídos na disciplina.

Serão elas:

- Atividade 1: 20,0 pontos
- Atividade 2: 20,0 pontos
- Atividade 3: 20,0 pontos
- Atividade 4: 20,0 pontos
- Atividade 5: 20,0 pontos

EXAME ESPECIAL: 23/07/2024

<b>CRONOGRAMA – CONTEÚDO E ATIVIDADES</b> (sujeito a mudanças e adaptações ao longo do semestre)	
<b>CONTEÚDO/ ATIVIDADES</b>	<b>SEMANAS</b>
<p>Definição de Mineração de Dados. Relação do processo de mineração de dados com descoberta de conhecimento, estatística, inteligência computacional. Fases de um processo de descoberta de conhecimento: obtenção e normalização de dados, limpeza de dados, seleção e transformação, mineração, avaliação do conhecimento.</p> <p>Leitura dos slides da disciplina</p> <p>Leitura dos capítulos 1, 2 e 3 do livro texto 1</p> <p>Leitura dos capítulos 1 e 2 do livro texto 2</p> <p><b>Trabalho prático 01 (22 de abril de 2024)</b></p>	<p>25 de março a 22 de abril (4 semanas)</p>
<p>Análise preditiva - Tarefas de mineração de dados: classificação e regressão.</p> <p>Leitura dos slides da disciplina</p> <p>Leitura dos capítulos 5 e 6 do livro texto 1</p> <p>Leitura do capítulo 3 do livro texto 2</p> <p><b>Trabalho prático 02 (13 de maio de 2024)</b></p>	<p>23 de abril a 13 de maio (3 semanas)</p>
<p>Análise de agrupamentos: A tarefa de agrupamento. Avaliação de modelos para análise de agrupamento.</p> <p>Leitura dos slides da disciplina</p>	<p>14 de maio a 03 de junho (3 semanas)</p>

<p>Leitura do capítulo 4 do livro texto 1</p> <p>Leitura do capítulo 4 do livro texto 2</p> <p><b>Trabalho prático 03 (03 de junho de 2024)</b></p>	
<p>Regras de associação: Processo de mineração de regras de Associação. Algoritmos para mineração de regras de associação.</p> <p>Leitura dos slides da disciplina</p> <p>Leitura do capítulo 7 do livro texto 1</p> <p>Leitura do capítulo 5 do livro texto 2</p> <p><b>Trabalho prático 04 (24 de junho de 2024)</b></p>	<p>04 de junho a 24 de junho (3 semanas)</p>
<p>Deteção de anomalias: Introdução. Processo de deteção de anomalias. Métodos estatísticos. Métodos algorítmicos.</p> <p>Leitura dos slides da disciplina</p> <p>Leitura do capítulo 8 do livro texto 1</p> <p><b>Trabalho prático 05 (15 de julho de 2024)</b></p>	<p>25 de junho a 15 de julho (3 semanas)</p>
<p><b>Exame Especial:</b></p>	<p><b>23 de julho de 2024</b></p>
<p>Bibliografia básica (LIVROS TEXTO)</p> <p>1. CASTRO, Leandro Nunes,; FERRARI, Daniel Gomes. <b>Introdução à Mineração de Dados: Conceitos básicos, algoritmos e aplicações</b>. São Paulo: Editora Saraiva, 2016. (consta na biblioteca online da UFOP, no sítio</p>	

[https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-](https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/cfi/0!/4/2@100:0.00)

[5/cfi/0!/4/2@100:0.00](https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/cfi/0!/4/2@100:0.00), acessado no dia 20/12/2020 às 17:00 hs. Este é o livro adotado para a disciplina e deve ser acessado exclusivamente para fins didáticos, sendo vedada sua utilização para quaisquer outros fins)

2. SILVA, Leandro Augusto; PERES, Sarajane Marques; BOSCARIOLI, Clodis. **Introdução à mineração de dados: com aplicações em R**. 1ª Ed. Rio de Janeiro: Elsevier, 2016 (consta na biblioteca online da UFOP, no sítio <https://integrada.minhabiblioteca.com.br/#/books/9788595155473/cfi/6/2!/4/2@0.00:0.00> acessado no dia 20/12/2020 às 17:10 hs. Este é o livro adotado para a disciplina e deve ser acessado exclusivamente para fins didáticos, sendo vedada sua utilização para quaisquer outros fins)

#### Bibliografia complementar

1. KUMAR, V. ; TAM P. Introdução ao Data Mining; Mineração de dados. Rio de Janeiro: Riachuelo, Ed. Ciência Moderna, 2009.
2. HAN, J. KAMBER, M. Data Mining.: Concepts and techniques. Morgana Kaufmann, 2000.
3. TAN, P.; STEIMBACH, M.; KUMATR, V. Introduction ao data mining. Boston, Addison Wesley, 2006
4. CARVALHO, L.A.V. Datamining A Mineração de Dados no Marketing, Medicina, Economia, Engenharia e
5. Administração. Rio de Janeiro: Editora Ciência Moderna, 2005.
6. T. Hastie, R. Tibshirani, J. Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd. Ed.) Springer, 2009.
7. GOLSCHIMIDT, R. Passos S. Data Mining; Um guia prático. Rio de Janeiro: Campus, 2009.
8. AMARAL, F.C.N. Data Mining: Técnicas e aplicações para o Marketing Direto. São Paulo: Ed. Berkeley, 2001.
9. BUSSAB, W.O. , MORETTIN, P.A. Estatística Básica. 5.ed. São Paulo: Editora Saraiva, 2002.
10. JOHNSON, R.A., WICHERN, D.W. Applied multivariate statistical analysis. 4 ed. New Jersey. Prentice Hall Inc. 1998.