



Universidade Federal
de São João del-Rei

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

PLANO DE ENSINO

Disciplina: Mineração de Dados Aplicada			Período: Variável	Currículo: 2017		
Docente: Leonardo Chaves Dutra da Rocha			Unidade Acadêmica: PPGCC			
Pré-requisito: -			Co-requisito: -			
C.H.	Total: 72ha/66h	Teórica: 54ha/49,5h	Prática: 18ha/16,5h	Grau: Mestrado	Ano: 2021	Semestre: PR01

Ementa

Processo de descoberta de conhecimento. Conceitos básicos de coleta e engenharia de dados. Técnicas de mineração de dados. Tarefas em mineração de dados. Domínios de aplicação.

Objetivos

Esta disciplina tem por objetivo apresentar várias das técnicas de mineração de dados, discutir detalhes da sua utilização e analisar campos de aplicações onde essas técnicas são utilizadas. Ao final dessa disciplina, o aluno deverá ser capaz de projetar e avaliar um processo de descoberta de conhecimento completo utilizando bases de dados reais.

Conteúdo Programático

- 1. Processo de descoberta de conhecimento**
 1. Preparação dos dados
 2. Mineração de dados
 3. Visualização dos resultados
- 2. Conceitos básicos de coleta e engenharia de dados**
 1. Exploração estática
 2. Tratamento dos valores ausentes
 3. Normalização
 4. Binning
 5. Seleção e Construção de características (Redução de Dimensão)
 6. Discretização
- 3. Técnicas de mineração de dados**
 1. Classificação
 2. Agrupamento
 3. Regras de Associação
- 4. Tarefas em mineração de dados**
 1. Avaliação objetiva da qualidade dos padrões
 2. Construção de padrões mais relevantes
 3. Visualização
- 5. Domínios de aplicação**
 1. Mineração de dados de aplicações Internet
 2. Comércio eletrônico
 3. Bioinformática

Metodologia de Ensino

O curso será desenvolvido mesclando aulas expositivas remotas síncronas e assíncronas, conforme planejamento apresentado em anexo. As aulas expositivas remotas síncronas serão ministradas por meio da aplicação GoogleMeet. As aulas assíncronas serão dedicadas à leitura de artigos para a elaboração do seminário. Ao final do semestre o aluno deverá realizar a apresentação do seminário e a apresentação de seu trabalho prático, conforma planejamento em anexo. Para ambos os casos, o aluno pode preparar um vídeo e disponibilizá-lo antes da data agendada para a sua apresentação. Todo material de aula será disponibilizado para o aluno por meio da plataforma CampusVirtual da UFSJ. As entregas previstas também se darão por meio do CampusVirtual da UFSJ.

Com relação ao horário de atendimento, o mesmo se dará nas sextas-feiras, de 13:00 às 16:00, mediante agendamento pelo e-mail lrocha@ufsj.edu.br com 24 horas de antecedência.

Controle de Frequência e Critérios de Avaliação

A avaliação do aprendizado será realizada através de um seminário (30%) e um trabalho prático o qual será dividido em 5 partes principais: 1) Proposta (15%), 2) Resultados Parciais 1 (10%), 3) Resultados Parciais 2 (10%), 4) Relatório (20%) e 5) Apresentação (15%). Além disso, será aplicada uma avaliação substitutiva, uma prova que versará sobre todo conteúdo, no valor de 30%. Terá direito à

realização dessa prova os alunos que obtiverem nota inferior a 60% da nota de todo semestre. Para o estudo, planejamento e execução do trabalho prático, estimamos em torno de 56 horas (4 horas semanais). Para o estudo, planejamento e preparação do seminário estimamos em torno de 14 horas (1 hora por semana).

A frequência dos alunos estará vinculada às etapas de entrega das atividades acima mencionadas, conforme apresentado abaixo e detalhado no planejamento em anexo (cada aula corresponde à duas horas/aula):

Aulas 1,2,3,4,5, → Estão vinculadas à proposta do trabalho
Aulas 6,7,8,9,10- → Estão vinculadas aos resultados parciais 1 do trabalho
Aulas 11,12,13,14,15, → Estão vinculadas aos resultados parciais 2 do trabalho
Aulas 16,17,18,19,20 → Estão vinculadas à entrega do relatório do seminário
Aulas 21,22,23,24,25,- → Estão vinculadas à apresentação do seminário
Aulas 26,27,28,29,30 → Estão vinculadas ao relatório do trabalho
Aulas 31,32,33,34,35,36 → Estão vinculadas a apresentação do trabalho

Bibliografia Básica

1. J. Han and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2nd Edition, 2007.
2. T AN, P.; STEINBACH, M.; KUMAR, V. Introdução ao Data Mining (Mineração de Dados) 2a Ed. Ciência Moderna, 2009
3. M. Zaki and W. Meira Jr. Fundamentals of Data Mining Algorithms, Cambridge, 2012
(a ser publicado, mas atualmente pode ser obtido por meio do endereço eletrônico:
<http://www.dcc.ufmg.br/miningalgorithms/DokuWiki/doku.php>)

Bibliografia Complementar

1. P. Tan, M. Steinbach, and V. Kumar Introduction to Data Mining Addison Wesley, 2006.
2. Documentação do Knime: <http://tech.knime.org/knime>
3. Artigos publicados nas principais conferências de mineração de dados:
KDD (<http://www.informatik.uni-trier.de/~ley/db/conf/kdd/index.html>),
SDM (<http://www.informatik.uni-trier.de/~%20ley/db/conf/sdm/index.html>),
SIGMOD (<http://www.informatik.uni-trier.de/~ley/db/conf/sigmod/index.html>),
VLDB (<http://www.informatik.uni-trier.de/~ley/db/conf/vldb/index.html>),
ICDM (<http://www.informatik.uni-trier.de/~ley/db/conf/icdm/index.html>), e
ICDE. (<http://www.informatik.uni-trier.de/~ley/db/conf/icde/index.html>)
4. - R. Elmasri, S. B. Navathe, Sistemas de Banco de Dados, Pearson, 2010.
- 5- A. Silberschatz, H. F. Korth, S. Sudarshan, Sistema de Bancos de Dados, Campus, 2006.



Leonardo Chaves Dutra da Rocha

Aprovado pelo Colegiado em

Coordenador

Planejamento de Aulas

Aula	Data	Conteúdo
1	18/05/2021	Conceitos básicos (Processo de Descoberta de Conhecimento)(Síncrona)
2	20/05/2021	Quatro técnicas tradicionais de Mineração de dados (Técnicas de mineração de dados) (Síncrona)
3	21/05/2021	Elaboração de Seminário (Assíncrona)
4	25/05/2021	Quatro técnicas tradicionais de Mineração de dados (Técnicas de mineração de dados)
5	27/05/2021	Preparando sua base – técnicas de tratamento de texto (Conceitos básicos de coleta e engenharia de dados) (Síncrona)
6	28/05/2021	Elaboração de Seminário (Assíncrona)
7	01/06/2021	Preparando sua base – técnicas de tratamento de texto (Conceitos básicos de coleta e engenharia de dados)(Síncrona)
8	03/06/2021	Elaboração de Seminário (Assíncrona)
9	08/06/2021	Pre-Processamento de dados: Exploração estática; Tratamento dos valores ausentes; Normalização; Binning Seleção e Construção de características; Discretização (Conceitos básicos de coleta e engenharia de dados) (Síncrona)
10	10/06/2021	Pre-Processamento de dados: Exploração estática; Tratamento dos valores ausentes; Normalização; Binning Seleção e Construção de características; Discretização (Conceitos básicos de coleta e engenharia de dados) (Síncrona)
11	11/06/2021	Elaboração de Seminário (Assíncrona)
12	15/06/2021	Transformação de valores numéricos (Conceitos básicos de coleta e engenharia de dados)(Síncrona)
13	17/06/2021	Transformação de valores numéricos (Conceitos básicos de coleta e engenharia de dados)(Síncrona)
14	18/06/2021	Elaboração de Seminário (Assíncrona)
15	22/06/2021	Apresentação da Proposta do Trabalho (Domínios de Aplicação) (Síncrona)
16	24/06/2021	Agrupamento (Técnicas de Mineração de Dados) (Síncrona)
17	25/06/2021	Elaboração de Seminário (Assíncrona)
18	29/06/2021	Agrupamento (Técnicas de Mineração de Dados) (Síncrona)
19	01/07/2021	Classificação (Técnicas de Mineração de Dados) (Síncrona)
20	02/07/2021	Elaboração de Seminário (Assíncrona)
21	06/07/2021	Classificação (Técnicas de Mineração de Dados) (Síncrona)
22	08/07/2021	Apresentação dos Resultados Parciais 1 do Trabalho (Domínios de Aplicação) (Síncrona)
23	09/07/2021	Elaboração de Seminário (Assíncrona)
24	13/07/2021	Regras de Associação (Técnicas de Mineração de Dados) (Síncrona)
25	15/07/2021	Regras de Associação (Técnicas de Mineração de Dados) (Síncrona)
26	16/07/2021	Elaboração de Seminário (Assíncrona)
27	20/07/2021	Mineração de Itemsets (Técnicas de Mineração de Dados) (Síncrona)
28	22/07/2021	Mineração de Itemsets (Técnicas de Mineração de Dados) (Síncrona)
29	27/07/2021	Apresentação dos Resultados Parciais 2 do Trabalho (Domínios de Aplicação) (Síncrona)
30	29/07/2021	Seminários (Tópicos Avançados em Mineração de dados) (Síncrona)
31	03/08/2021	Seminários (Tópicos Avançados em Mineração de dados) (Síncrona)
32	05/08/2021	Seminários (Tópicos Avançados em Mineração de dados) (Síncrona)
33	10/08/2021	Seminários (Tópicos Avançados em Mineração de dados) (Síncrona)
34	12/08/2021	Apresentação TP (Domínios de Aplicação) (Síncrona)
35	17/08/2021	Apresentação TP (Domínios de Aplicação) (Síncrona)
36	19/08/2021	Apresentação TP (Domínios de Aplicação) (Síncrona)
Aulas 1,2,3,4,5 → Estão vinculadas à proposta do trabalho		
Aulas 6,7,8,9,10 → Estão vinculadas aos relatórios parciais 1 do trabalho		
Aulas 11,12,13,14,15, → Estão vinculadas aos resultados parciais 2 do trabalho		
Aulas 16,17,18,19,20 → Estão vinculadas à entrega do relatório do seminário		
Aulas 21,22,23,24,25,- → Estão vinculadas à apresentação do seminário (por vídeo ou em aulas síncronas)		
Aulas 26,27,28,29,30 → Estão vinculadas ao relatório do trabalho		
Aulas 31,32,33,34,35,36 → Estão vinculadas a apresentação do trabalho (por vídeo ou aulas síncronas)		
Os seminários e as apresentações dos trabalhos serão síncronas, porém os alunos poderão compartilhar um vídeo		