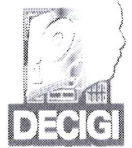


PLANO DE ENSINO DE DISCIPLINA
(FICHA Nº 2)

Disciplina: Mineração de Dados	Código: SIN134
Válido para: 2016 (X) 1º Semestre () 2º Semestre Anual () Modular ()	
Ofertada para o curso: Gestão da Informação	

DADOS DA FICHA 1	Natureza: (X) obrigatória () optativa
	Modalidade: (X) Presencial () EaD () 20% EaD
	Pré-requisito(s): SIN133 – Banco de Dados
	Co-requisito(s): Não há
	C.H. Total: 60 C.H. Semanal: Padrão: 02 Laboratório: 02 Campo: 00 Estágio: 00 Orientação: 00



EMENTA

Técnicas de descoberta de conhecimento em bancos de dados utilizando as diversas tarefas de mineração de dados.

UNIDADES DIDÁTICAS

1. Apresentação (2 horas)
 - 1.1. Posicionamento
 - 1.2. Motivação
 - 1.3. Conceitos Básicos (IA, DW)
2. KDD e mineração de dados (6 horas)
 - 2.1. Aplicações
 - 2.2. Etapas do processo de KDD
 - 2.3. Mineração de dados
 - 2.4. Técnicas e tarefas de mineração de dados
 - 2.5. Outros: base de dados; conhecimento; transformação; atributos; partição e árvore de decisão.
3. Preparando os dados para mineração de dados (2 horas)
 - 3.1. Dados necessários em mineração de dados
 - 3.2. Estrutura ideal dos dados para mineração de dados
4. Técnicas, heurísticas, algoritmos (2 horas)
 - 4.1. Noções gerais
 - 4.2. Heurísticas diversas
 - 4.2.1. Redes neurais
 - 4.2.2. Lógica nebulosa
 - 4.2.3. Algoritmos genéticos
5. Tarefas (24 horas)
 - 5.1. Associação
 - 5.2. Agrupamento
 - 5.3. Classificação
6. Tarefas (24 horas)
 - 6.1. Associação
 - 6.2. Agrupamento
 - 6.3. Classificação
7. Pós-processamento (4 horas)
 - 7.1. Simplificações
 - 7.2. Formatos de apresentação de resultados
 - 7.3. Avaliação de resultados
8. Aplicações dos algoritmos às bases de dados (20 horas)
 - 8.1. Aplicativo Weka
 - 8.2. Outros aplicativos (R, Rapidminer etc)

OBJETIVOS

O aluno deverá conhecer e ser capaz de aplicar diferentes técnicas estatísticas e computacionais de exploração de dados na identificação de padrões relevantes e não óbvios em bancos de dados.

O aluno deverá ser capaz de:

- 1 discutir os conceitos básicos relacionados à mineração de dados;
- 2 aplicar e explicar resultados das principais tarefas de mineração de dados;
- 3 justificar a relevância desta disciplina;
- 4 identificar e compreender aplicações práticas da mineração de dados.

PROCEDIMENTOS DIDÁTICOS

A disciplina consiste de duas partes distintas:

- 1- Informação Teórica: teoria e algoritmos de descoberta de conhecimento em banco de dados. Itens 1 a 7. No item 5 - Tarefas, estão previstas 6 (seis) aulas na modalidade à distância;
- 2- Análise de algoritmos em laboratório: utilização e interpretação da utilização de algoritmos de mineração de dados em laboratório de informática. Item 8.

A primeira parte será apresentada de forma expositiva, seguidas de tarefas / exercícios de aplicação (conceituais) que devem ser entregue ao final da aula ou no início da aula seguinte (conforme especificado).

A segunda parte, a partir do item 8 será composta de seminários a serem preparados e apresentados pelos próprios alunos sob supervisão dos responsáveis pela disciplina.

TODAS AS AVALIAÇÕES (PROVAS) SERÃO INDIVIDUAIS e PRESENCIAIS.

RECURSOS

Quadro-de-giz, multimídia, sala virtual no ambiente MOODLE

FORMAS DE AVALIAÇÃO

As três notas principais para composição da média final (cálculo da média aritmética simples das três notas), sendo:

- a) primeira nota: média aritmética das atividades realizadas em sala de aula ou propostos em aula e feitos individualmente ou em equipe (**conforme solicitado**);
- b) segunda nota: média aritmética das notas das provas escritas (**individuais**), sobre o conteúdo da parte 1;
- c) terceira nota: trabalho em dupla / individual sobre um dos temas disponíveis na segunda parte da disciplina (item 7), devendo a avaliação considerar:
 - profundidade na abordagem do tema;
 - qualidade do levantamento bibliográfico;
 - contextualização do tema na área de interesse;
 - explicação do funcionamento do algoritmo em questão;
 - exemplificações práticas e contextualizadas do tema;
 - análises dos resultados obtidos;
 - criatividade da apresentação;
 - uso de normas para elaboração de trabalhos de pesquisa (normas ABNT);
 - desempenho individual e coletivo dos membros da equipe durante a apresentação.

Obrigatória frequência mínima de 75% às aulas



BIBLIOGRAFIA BÁSICA

CARVALHO, L. A. V. de. **Datamining**: a mineração de dados no marketing, medicina, economia, engenharia e administração. São Paulo: Érica, 2002. (5 ex)

DAUMÉ III, Hal. **A course in machine learning**. Disponível em: <<http://ciml.info/>>. Acesso em: 21 jul. 2015.

Machine learning: the complete guide. Disponível em: <https://en.wikipedia.org/wiki/Book:Machine_Learning_%E2%80%93_The_Complete_Guide>. Acesso em 21 jul. 2015.

Data mining algorithms in R. Disponível em: <https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R>. Acesso em 21 jul. 2015.

BIBLIOGRAFIA COMPLEMENTAR

AMARAL, F.C.N. **Data mining**: técnicas e aplicações para o marketing direto. São Paulo: Editora Berkeley, 2001.

BUSSAB, W.O. , MORETTIN, P.A. **Estatística básica**. 5.ed. São Paulo: Editora Saraiva, 2002.

GOLDSCHMIDT, R. R. ; PASSOS, E. P. L. **Data mining**: um guia prático - conceitos, técnicas, ferramentas, orientações e aplicações. 1. ed. Rio de Janeiro: Editora Campus, 2005. v. 1. 250 p.

BENOIT, G. **Data mining**. In: **Annual Review of Information Science and Technology**, v. 36, nº 1, p. 265–310, 2002. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/aris.1440360107/pdf>>. Acesso em: 24 jul. 2015.

WEKA. **Waikato Environment for Knowledge Analysis**. The University of Waikato. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 24 jul. 2015.

Curitiba, 04 de dezembro de 2015.

Professor da Disciplina: Denise Fukumi Tsunoda
Assinatura:

Coordenação do Curso: Rodrigo Eduardo Botelho Francisco
Assinatura:

Prof. Dr. Rodrigo Botelho Francisco
Departamento de Ciência e gestão
da Informação
Matrícula: 1422358

Chefe de Departamento: Edmeire Cristina Pereira
Assinatura: