



## PLANO DE ENSINO

Ficha nº 2

Disciplina: Tópicos em Informática Aplicada II

Código: SIN090

Departamento/Instituto: Ciência e Gestão da Informação

Curso: Gestão da Informação

Válido para o ano de: 2010

( X ) 1º semestre ( ) 2º semestre

Carga horária semanal: 04 horas

Créditos: 03

Aulas: 30 teóricas 30 práticas

Total: 60 horas

Turma(s): A

Horário(s): Terça-feira: 10:30 – 12:30

Quarta-feira: 10:30 – 12:30

### **OBJETIVOS GERAIS DA DISCIPLINA:**

Apresentar e aplicar diferentes técnicas estatísticas e computacionais de exploração de dados para identificar padrões importantes e não óbvios em grandes bancos de dados

### **EMENTA:**

Uso de ferramentas e aplicativos da Informática para gestão da informação.

### **CONTEÚDO PROGRAMÁTICO:**

1. Apresentação (2 horas)

1.1. Posicionamento

1.2. Motivação

1.3. Conceitos Básicos

2. KDD e mineração de dados (6 horas)

2.1. Observando e aprendendo

2.2. Etapas do processo de KDD

2.3. A metodologia de data mining

2.4. As técnicas e tarefas de data mining

2.5. Outros: base de dados; conhecimento; transformação; atributos; partição; substituição; regressão; score; árvore de decisão e relatório.

3. Preparando os dados para data mining (4 horas)

3.1. Dados necessários em data mining

3.2. Estrutura ideal dos dados para data mining

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 (12 horas)

5.1. Associação

5.2. Agrupamento



- 5.3. Classificação
- 6. Pós-processamento (2 horas)
  - 6.1. Simplificações
  - 6.2. Visualização
  - 6.3. Avaliação de resultados
- 7. Aplicações dos algoritmos às bases de dados (30 horas)
  - 7.1. Aplicativo Weka
  - 7.2. Outros aplicativos
  - 7.3. Aplicação dos algoritmos
  - 7.4. Trabalhos

### **METODOLOGIA:**

A disciplina consiste de duas partes distintas:

- 1- Informação Teórica: teoria e algoritmos de descoberta de conhecimento em banco de dados. Itens 1 a 6;
- 2- Análise de algoritmos em laboratório: utilização e interpretação da utilização de algoritmos de mineração de dados em laboratório de informática. Item 7.

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 7 será composta de seminários a serem preparados e apresentados pelos próprios alunos sob supervisão dos responsáveis pela disciplina.

Algumas aulas (não ultrapassando o limite de 20% do total de horas da disciplina) serão ministradas na modalidade à distância. Ressalta-se que todas as avaliações serão realizadas na forma PRESENCIAL, seguindo normatizações para desenvolvimento de atividades didáticas a distância na Portaria 4.059 de 10 de dezembro de 2004 do Ministério da Educação.

### **FORMAS DE AVALIAÇÃO:**

Duas notas principais para composição da média final (média aritmética das duas notas principais), sendo:

Primeira nota: média aritmética de: nota da prova escrita (**individual** e **sem consulta**), sobre o conteúdo da parte 1 + média dos exercícios propostos em aula e feitos individualmente ou em equipe (conforme solicitado);

Segunda nota: trabalho individual ou em dupla 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




## BIBLIOGRAFIA:

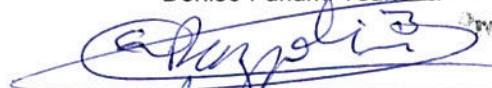
- AMARAL, F.C.N. **Data Mining: Técnicas e Aplicações para o Marketing Direto**. São Paulo: Editora Berkeley, 2004.
- BUSSAB, W.O., MORETTIN, P.A. **Estatística Básica**. 5.ed. São Paulo: Editora Saraiva, 2002.
- BUSSAB, W.O., MIAZAKI, E.S. ANDRADE, D.F. **Introdução à Análise de Agrupamentos**. São Paulo: 9º Simpósio Nacional de Probabilidade e Estatística, 1990.
- BERRY, M.J.A., LINOFF, G. **Data Mining Techniques for Marketing, Sales and Customer Support**. 2ª. ed. New York: John Wiley & Sons, Inc., 2004.
- CARVALHO, L.A.V. **Datamining – a mineração de dados no marketing, medicina, economia, engenharia e administração**. São Paulo: Editora Érica, 2001.
- DINIZ, C.A.R., NETOF, L. **Data Mining: uma introdução**. São Paulo: XIV Simpósio Nacional de Probabilidade e Estatística. IME-USP, 2000.
- FERNANDEZ, G. **Data Mining Using SAS Applications**. New York: Editora Chapman & Hall/CRC, 2003.
- 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.
- HAN, J., KAMBER, M. **Data Mining: Concepts and Techniques**. San Francisco: Morgan Kaufmann Publishers, 2001.
- KLÖSGEN, W., ZYTKOW, J.M. **Handbook of Data Mining and Knowledge Discovery**. New York: Oxford University Press, Inc., 2002.
- MANLY, B.F.J. **Multivariate Statistical Methods: a primer**. 3.ed. New York: Chapman & Hall, 2005.
- TAN, P. N., Steinbach, M. e Kumar, V. **Introduction to Data Mining**. Addison Wesley, 2005.
- WEKA. **Waikato Environment for Knowledge Analysis**. The University of Waikato. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 02 fev. 2010.

## ASSINATURAS:

Professora Responsável

  
Denise Fukumj Tsunoda

Coordenador do Curso

  
Prof. Dr. Edelvino Razzolini Filho  
Coordenador do Curso de  
Gestão da Informação - UFPR  
Matricula 200.114

Chefe do Departamento

  
Mauro José Belli  
Prof. Dr. Mauro José Belli  
Chefe do Departamento de Ciência  
Gestão da Informação - UFPR  
Matr. 120294