



Universidade Federal do Espírito Santo
Centro de Ciências Jurídicas e Econômicas
Departamento de Arquivologia

Plano de Ensino

Universidade	: Universidade Federal do Espírito Santo
Campus	: Goiabeiras
Curso	: Arquivologia - Noturno
Departamento Responsável	: Departamento de Arquivologia - CCJE
Data de Aprovação (Art. nº 91)	:
Docente Principal	: Jean-Rémi Bourguet
Matrícula	: 3039166
Qualificação	: http://lattes.cnpq.br/2469487264010076
Disciplina	: Ciências de Dados
Código	: ARV12967
Período	: 2019/1
Turma	: 1
Pré-requisito	: ARV13049
Carga Horária Semestral (CHS)	: 60hs
Distribuição da CHS	: Teórica:30hs / Exercício:15hs / Laboratório:15hs
Créditos	: 2

EMENTA

Metodologias e técnicas da ciência de dados para armazenamento, análise, tratamento, acessibilidade e visualização de dados.

OBJETIVOS ESPECÍFICOS

O aluno deverá conhecer as metodologias e técnicas da ciência de dados para armazenamento, análise, tratamento, acessibilidade e visualização de dados.

CONTEÚDO PROGRAMÁTICO

I. Coleta de Dados

- a. Fonte de Dados**
- b. Limpeza de Dados**

II. Mineração de Dados

- a. Análise Estatística**
- b. Aprendizagem**

III. Visualização de Dados

- a. Codificação Visual**
- b. Análise Visual**

METODOLOGIA

Será utilizada uma sistemática de aulas teóricas expositivas dialogadas e auxílio de quadro branco e projetor multimídia. Uma parte das atividades será realizada em laboratório na forma de Trabalhos Práticos. O material criado para a disciplina estará disponível no AVA (Ambiente Virtual de Aprendizagem) da UFES, em <http://ava.ufes.br> → Departamento de Arquivologia → Ciências de Dados (Prof. Jean-Rémi Bourguet). Vale ressaltar que algumas atividades serão executadas à distância, no AVA.

PROCESSO DE AVALIAÇÃO DA APRENDIZAGEM

O aluno será avaliado através de exames; projetos individuais ou em grupo realizados no horário da aula ou fora delas; atividades online e/ou presenciais. Todas essas avaliações serão anunciadas ao longo do curso. As datas dos exames serão combinadas com os alunos no decorrer do período. A média parcial (MP) será calculado da seguinte forma:

$$MP = \frac{10 \times \sum_i T_i}{100} + \frac{30 \times \sum_j P_j}{100} + \frac{60 \times \sum_k E_k}{100}$$

Onde T_i é uma atividade avulsa dada em sala de aula ou através de meio eletrônico, P_j é um projeto que poderá ser desenvolvido ao longo da disciplina e, por fim, E_k é um exame (prova) a ser realizado sobre o conteúdo da disciplina. A soma das notas máximas dos T_i s, P_j s e E_k s são cada uma iguais a 10.

- Caso MP seja maior ou igual a 7,0, o aluno estará aprovado.
- Caso MP seja menor que 7,0, o aluno terá oportunidade de fazer a prova final.

Neste último caso, a nota da prova final (NPF) terá pontuação de 0 a 10, e a média final (MF) será composta pela média aritmética entre a média parcial e a nota da prova final:

$$MF = \frac{MP + NPF}{2}$$

BIBLIOGRAFIA BÁSICA

CARVALHO, L. A. V. Data Mining: A mineração de dados no marketing, medicina, engenharia e administração. São Paulo: Érica, 2001.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMITH, P. Knowledge Discovery and Data Mining: Towards a Unifying Framework. In Proceedings of the Second International Conference on Data Mining and Knowledge Discovery, AAAI Press, Menlo Park, US; 1996. Disponível em <https://www.aaai.org/Papers/KDD/1996/KDD96-014.pdf>. Acesso em Maio de 2016.

REZENDE, Solange Oliveira. Mineração de Dados. Anais do V Encontro Nacional de Inteligência Artificial, Sociedade Brasileira de Computação. ISBN: 8576690330. São Leopoldo RS, 25 a 29 de julho de 2005. Disponível em <http://www.lbd.dcc.ufmg.br/colecoes/enia/2005/0102.pdf>. Acesso em Maio de 2016.

SILVA, Leandro. A.; SILVA Luciano. Fundamentos de Mineração de Dados Educacionais. Anais dos Workshops do Congresso Brasileiro de Informática na Educação. 3º Congresso Brasileiro de Informática na Educação (CBIE 2014) Workshops (WCBIE 2014). Disponível em <http://dx.doi.org/10.5753/cbie.wcbie.2014.568>. Acesso em Maio de 2016.

BIBLIOGRAFIA COMPLEMENTAR

BAKER, R.; ISOTANI, S.; DE CARVALHO, A. Mineração de Dados Educacionais: Oportunidades para o Brasil. Revista Brasileira de Informática na Educação, Volume 19, Número 2, 2011. Disponível em <http://www.columbia.edu/~rsb2162/BD-RBIE-pt-v22.pdf>. Acesso em Maio de 2016.

CHEN, Hsinchun; CHIANG, Roger HL; STOREY, Veda C. Business Intelligence and Analytics: From Big Data to Big Impact. MIS quarterly, v. 36, n. 4, p. 1165-1188, 2012. Disponível em <http://hmchen.shidler.hawaii.edu/Chen%5fbig%5fdata%5fMISQ%5f2012.pdf>. Acessado em Junho de 2015.

FERRARI, Luciana Itida. Uma Metodologia para Extração de Informação sobre o Sistema Imunológico. Tese (Doutorado). Coordenação do Programa de pós-graduação em Engenharia de Sistemas e Computação, Universidade Federal do Rio de Janeiro (COPPE/UFRJ), RJ, 2008.

PIATETSKY-SHAPIRO, G. FAYYAD, U. An Introduction to SIGKDD and A Reflection on the Term 'Data Mining'. SIGKDD Explorations Volume 13, Issue 2, Page 102, 2011. Disponível em <http://www.kdd.org/explorations/files/V13-02-24-Piatetsky-Shapiro.pdf>. Acesso em Maio de 2016.