



PLANO DE ENSINO

1. DISCIPLINA:

Nome da Disciplina: Mineração de Dados e Descoberta de Conhecimento		Código da Disciplina:
Professor Responsável: Jerry Adriani Johann		
Programa: Pós-graduação em Engenharia Agrícola – PGEAGRI		
Área de Concentração: Sistemas Biológicos e Agroindustriais & Recursos Hídricos e Saneamento Ambiental		
Centro: Ciências Exatas e Tecnológicas - CCET		
Campus: Cascavel		
Nível: Mestrado () Mestrado e Doutorado (X)	Doutorado ()	Semestre de oferta: Segundo Ano de oferta: 2020
Carga horária total: 60 h/a		Carga horária teórica: Carga horária de aulas prática:

2. EMENTA:

Conceitos de descoberta de conhecimento em banco de dados; conceitos e aplicações de mineração de dados (*Data Mining*); caracterização e representação de dados; métodos de preparação de dados para mineração; principais tarefas de mineração de dados (classificação, regressão, sumarização, regras de associação, agrupamento, anomalias); extração de padrões; pós-processamento; estudos de casos em agricultura.

3. OBJETIVOS:

Subsidiar os interessados na descoberta automática de informações úteis, novas e compreensíveis em bases de dados, combinando técnicas tradicionais de análise de dados com algoritmos sofisticados com o propósito de transformar dados, aparentemente ocultos, em informações úteis para a tomada de decisão e/ou avaliação de resultados.

4. CONTEÚDO PROGRAMÁTICO:

1. Introdução a Mineração de Dados (*Data Mining*).
2. Conceitos de Descoberta de Conhecimento em Banco de Dados.
3. Preparação de dados para Mineração de Dados (*Data Mining*):
 - 3.1 Caracterização, Representação e Qualidade de Dados.
 - 3.2 Métodos de pré-processamento de Dados: limpeza dos dados; extração e integração; seleção e transformação; redução de dimensionalidade.
4. Tarefas de Mineração de Dados:
 - 4.1. Preditivas: classificação e regressão.
 - 4.2. Descritivas: associação, agrupamento, sumarização.
5. Pós-processamento: avaliação de padrões.
6. Estudos de casos em Agricultura.

5. ATIVIDADES PRÁTICAS (grupo de ____ alunos):

6. METODOLOGIA:

Aulas teóricas e práticas. Resolução de exercícios. Uso de ferramentas computacionais em sala de aula ou de forma remota em que cada em que cada aluno matriculado fará uso de seu computador pessoal. Discussão de artigos técnicos científicos. Uso de ambiente virtual de aprendizagem.
Considerando a Resolução nº 052/2020-CEPE, no período de suspensão das atividades acadêmicas letivas presenciais, a disciplina será ministrada por aulas remotas síncronas nas Plataformas Microsoft Teams ou Google Meet. Será criada uma Sala Virtual da disciplina e todos os alunos matriculados serão inseridos nela para acompanhamento das aulas remotas. O controle de frequência dos alunos será realizado por meio da verificação dos alunos presentes nas salas virtuais de ensino, enquanto as atividades forem de forma remota. Caso no decorrer da disciplina, seja aprovado pelas instâncias superiores da universidade, o retorno as atividades presenciais, as aulas deixarão de ser remotas e passarão a serem presenciais.



7. AVALIAÇÃO (critérios, mecanismos, instrumentos e periodicidade):

A nota final (NF) da disciplina (Equação 1) será a média ponderada de cinco atividades (A, B, C, D, E).

$$NF = (A * 0,10) + (B * 0,25) + (C * 0,20) + (D * 0,20) + (E * 0,25) \quad (1)$$

em que: NF = Nota final; A = Revisão de artigos científicos; B = Listas de exercícios; C = Seminários; D = Trabalho de Conclusão da Análise de um Banco de Dados (apresentação oral e versão escrita do artigo); E = Artigo técnico-científico final do trabalho de conclusão

8. BIBLIOGRAFIA BÁSICA:

- BARI, A.; CHAUNCHI, M.; JUNG, T. **Análise Preditiva: para leigos**. Rio de Janeiro: Atlas Books, 464p. 2019.
- FAYYAD, U.; SHAPIRO, G. P.; SMYTH, P. Knowledge discovery and data mining: towards a unifying framework. In: **Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)**, 1996.
- GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. **Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações**. RJ: Campus, 296p. 2015.
- TAN, P-N; STEINBACH, M; KUMAR, V. **Introdução ao Data Mining: Mineração de Dados**. SP: Ciência Moderna, 900p., 2009.
- REZENDE, S. O. **Sistemas inteligentes. Fundamentos e Aplicações**. Barueri, SP: Manole, 525p. 2005.
- WEKA - Machine Learning Group at the University of Waikato. **Software**. 2009. Disponível em: <http://www.cs.waikato.ac.nz/ml/weka>. Consultado em 15 de fevereiro de 2009.
- WITTEN, I. H.; FRANK, E. **Data mining: Practical machine learning tools and techniques**. 2ª ed. São Francisco: Morgan Kaufmann, 525p., 2005.

9. BIBLIOGRAFIA COMPLEMENTAR:

- AMO, S. de. **Curso de Data Mining: Compilação de Notas de Aulas**, UFU, 201p., 2003.
- CEPEA - Centro de Estudos Avançados em Economia Aplicada. **Dados Agropecuários**. 2014. Disponível em: <http://www.cepea.esalq.usp.br>. Consultado em 20 de agosto de 2014.
- CONAB - Companhia Nacional de Abastecimento. **Dados Agropecuários**. 2014. Disponível em: <http://www.conab.gov.br>. Consultado em 20 de agosto de 2013.
- FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. **From Data Mining to Knowledge Discovery in Databases**. American Association for Artificial Intelligence, 1996.
- FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R., **Advances in Knowledge Discovery and Data Mining**, MIT Press, 1996.
- FAOSTAT - Food and Agriculture Organization of the United Nations. **Production and Crops in the World**. 2014. Disponível em: <http://faostat.fao.org/site/291/default.aspx>. Consultado em 20 de agosto de 2014.
- HAN, J. & KAMBER, M. **Data Mining: Concepts and Techniques**. 2nd Edition, Morgan Kaufmann, 746p. 2006.
- IBGE - Instituto Brasileiro de Geografia e Estatística. **Banco de Dados Agregados - Sistema IBGE de Recuperação Automática - SIDRA**. 2014. Disponível em: <http://www.sidra.ibge.gov.br>. Consultado em 20 de agosto de 2014.
- LAROSE, D.T. **Data Mining: Methods and Models**. Wiley-Interscience, 322p. 2006.
- MAIMON, O.; ROKACH, L. **Data Mining and Knowledge Discovery Handbook**. 2nd Edition, Springer, 1285p. 2010.
- OLSON, D.L.; DELEN, D. **Advanced Data Mining Techniques**. Springer, 182p. 2008.
- USDA - United States Department of Agriculture. **Foreign Agricultural Service**. 2014. Disponível em: <http://www.fas.usda.gov/data>. Consultado em 20 de agosto de 2014.

10. ASSINATURAS:

Assinatura do professor responsável:	Cascavel, 18 de setembro de 2020.
Assinatura e carimbo do coordenador do PGEAGRI: <i>[Assinatura]</i> RG 31724589 Coordenador do Programa de Pós-Graduação em Engenharia Agrícola	(X) Aprovado Ata Nº <u>04</u> do dia <u>23/09/2020</u>
Assinatura e carimbo do Diretor do PGEAGRI: <i>[Assinatura]</i> Engenharia Agrícola	(X) Homologado Ata Nº <u>04</u> do dia <u>30/09/2020</u>
Encaminhado cópia à secretaria acadêmica em:	<i>[Assinatura]</i>