



Plano de Disciplina

Identificação				
CURSO Mestrado Profissional em Tecnologia da Informação			CAMPUS João Pessoa	
DISCIPLINA Ciência de Dados			CÓDIGO DA DISCIPLINA 7791	
PRÉ-REQUISITO Não há pré-requisito				
UNIDADE CURRICULAR (OBRIGATÓRIA, OPTATIVA, ELETIVA) Eletiva			SEMESTRE 2025.1	
DOCENTES RESPONSÁVEIS Damires Yluska de Souza Fernandes Alex Sandro da Cunha Rego				
Carga Horária				
TEÓRICA 24	PRÁTICA 24	EaD ¹ 12	CARGA HORÁRIA SEMANAL: 4	CARGA HORÁRIA TOTAL 60

Ementa

Fundamentos de Ciência de Dados; Modelos de processos para descoberta de conhecimento; Técnicas de coleta, seleção, preparação, limpeza e pré-processamento de dados; Tipos de análises de dados; Identificação e tratamento de outliers; Visualização de dados; Modelagem de problemas de Aprendizado de Máquina com aplicação de algoritmo preditivo; Interpretação de resultados.

Objetivos

Conhecer conceitos, demandas e ferramentas relacionados à ciência de dados em diversos cenários e aplicações; identificar desafios para pesquisa e inovação em ciência de dados por meio da discussão de problemas em aberto e possibilidades de soluções para problemas práticos advindos da indústria e da academia em cenários atuais; Aplicar conceitos e ferramentas estudados em projeto prático em ciência de dados a partir de um problema/cenário real.

Conteúdo Programático

Tópico	Carga Horária
1. Fundamentos de Ciência de Dados	4
Introdução	2
Aplicações de Ciência de Dados	2
2. Modelos de Processo para Descoberta de Conhecimento	4
O processo KDD (<i>Knowledge Discovery in Databases</i>)	1
CRISP-DM (<i>Cross Industry Standard Process for Data Mining</i>)	2
Pipeline em Ciência de Dados	1

¹ Para a oferta de disciplinas na modalidade à distância, integral ou parcial, desde que não ultrapasse 20% (vinte por cento) da carga horária total do curso, observar o cumprimento da Portaria MEC nº 1.134, de 10 de outubro de 2016.



3. Coleta e Integração de Dados	8
Seleção de fontes de dados e métodos de coleta	2
Exploração e entendimento dos dados	2
Integração de Dados	2
Qualidade de Dados	2
4. Pré-processamento	10
Normalização	1
Padronização e conversão de tipos de dados	1
Discretização	1
Tratamento de valores ausentes	1
Remoção de duplicatas	1
Derivação de atributos	1
Tratamento de outliers (em nível de valor)	2
One-hot encoding	2
5. Análise de Dados	12
5.1 Análise baseada em consultas	1
5.2 Análise exploratória: descritiva, tendências, visual	3
5.2 Análise preditiva	8
Aprendizado Supervisionado: classificação	2
Algoritmo de Árvore de Decisão	3
Técnicas de avaliação de modelos	1
Métricas de avaliação	2
6. Desafios e Oportunidades	8
Seminários sobre temas em ciência de dados para pesquisas	8
7. Produção de conjuntos de dados para aplicação	14
Leitura de artigos, preparação do conjunto de dados	6
Submissão de artigo e apresentação de resultados	8

Metodologia de Ensino

Aulas teóricas e práticas.

Análises e discussões sobre cases, pesquisas e soluções no escopo da disciplina.

Planejamento e execução de projeto de aplicação em ciência de dados, visando à produção de conjunto de dados e sua aplicabilidade.

Reuniões de acompanhamento dos projetos para cada time de desenvolvimento e entregas em datas pré-estabelecidas.



Recursos Didáticos

Ferramentas de software open source, periódicos, artigos, classroom, laboratórios, acesso à Internet.

Critérios de Avaliação

A avaliação dar-se-á com base em **dois** marcos: (i) **Atividades** ao longo da disciplina (pesquisas, exercícios, seminários, apresentações) e (ii) **Projeto** e implementação de aplicação em ciência de dados, visando à produção de conjunto de dados e sua aplicabilidade, com demonstração, apresentação oral do trabalho, disponibilização do dataset em repositório e submissão de artigo. Todas as atividades terão prazos definidos a serem cumpridos.

Bibliografia

Básica

HAN, J.; KAMBER, M.; PEI, J. Data Mining: Concepts and Techniques. 3. ed. Waltham (USA): Morgan Kaufmann, 2012.

DE CASTRO, Leandro Nunes; FERRARI, Daniel Gomes. Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações. São Paulo: Saraiva, 2016.

Complementar

AMARAL, F. Introdução à Ciência de Dados. Rio de Janeiro: Alta Books, 2016.

VANDERPLAS, J. Python Data Science Handbook: Essential Tools for Working with Data. 1. ed. Sebastopol: O'Reilly Media, Inc., 2016.

GRUS, J. Data Science from Scratch: First Principles with Python. Sebastopol: O'Reilly Media, 2015.

ZAKI, M.; MEIRA, Jr., W. Data Mining and Machine Learning: Fundamental Concepts and Algorithms. 2. ed. Cambridge: Cambridge University Press, 2020. ISBN 978-1108473989.

ALPAYDIN, Ethem. Introduction to Machine Learning. 2. ed. Cambridge: MIT Press, 2010.

MITCHELL, T. Machine Learning. Nova Iorque: McGraw-Hill, 1997.

WITTEN, I.; FRANK, E. Data Mining: Practical Machine Learning Tools and Techniques. 3. ed. San Francisco: Morgan Kaufmann Publishers, 2011.

KELLEHER, J. D.; MAC NAMEE, B.; D'ARCY, A. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. Cambridge: MIT Press, 2015.

BÖRNER, K.; POLLEY, D. Visual Insights: A Practical Guide to Making Sense of Data. Cambridge, Mass.: The MIT Press, 2014. TELEA, A. Data Visualization: Principles and Practice. 2. ed. Boca Raton: A K Peters/CRC Press, 2014.

Observações

(Nenhuma)