



Plano de Disciplina

Identificação				
CURSO Mestrado Profissional em Tecnologia da Informação			CAMPUS João Pessoa	
DISCIPLINA Mineração de Dados			CÓDIGO DA DISCIPLINA 30630	
PRÉ-REQUISITO Conhecimentos sobre Banco de Dados (modelo relacional, SQL, projeto, transações, noções sobre modelos NoSQL e gerenciamento de dados em ambientes diversos).				
UNIDADE CURRICULAR (OBRIGATÓRIA, OPTATIVA, ELETIVA) Eletiva			SEMESTRE 2019.2	
DOCENTES RESPONSÁVEIS Damires Yluska de Souza Fernandes				
Carga Horária				
TEÓRICA 24	PRÁTICA 24	EaD ¹ 12	CARGA HORÁRIA SEMANAL: 4	CARGA HORÁRIA TOTAL 60

Ementa

Introdução à Mineração de Dados e conceitos básicos; Cenários de aplicação de mineração de dados; Processo KDD (*Knowledge Discovery in Databases*) e suas etapas; Tarefas de Mineração de Dados: Regras de Associação, Classificação, Regressão e Clusterização; Algoritmos de Mineração de Dados; Seleção de Atributos; Aprendizado supervisionado e não-supervisionado; Técnicas de avaliação; Avaliação, análise e interpretação de resultados.

Objetivos

Estudar e discutir conceitos, técnicas e soluções associadas a processos e aplicações com mineração de dados no contexto de big data. A motivação para a pesquisa e estudos nesta área é o desafio para lidar eficientemente com volumes massivos de dados heterogêneos produzidos por pessoas, dispositivos e sistemas em diversos domínios de aplicação de modo a prover análises e visualizações adequadas e que agreguem valor.

Objetivos específicos:

- Caracterizar e entender Big Data e Mineração de Dados;
- Entender as principais problemáticas associadas à gerência de Big Data e mineração;
- Estudar e aplicar as etapas necessárias para o desenvolvimento e uso de big data e mineração de dados;
- Estudar e aplicar técnicas de mineração de dados no contexto big data;
- Estudar e aplicar técnicas de visualização da informação pós mineração de dados;
- Pesquisar, sumarizar e apresentar temática relevante de pesquisa na área de mineração de dados e big data.

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.



Conteúdo Programático - Presencial

Tópico	Carga horária
Introdução à Mineração de Dados e conceitos afins	4
Mineração de dados e KDD na era Big Data: problemas, processos e aplicações	4
Processos de Integração de Dados	8
Tarefas de Mineração de Dados e aplicações	8
Aprendizado supervisionado: técnicas, algoritmos e avaliações	8
Aprendizado não supervisionado: técnicas, algoritmos e avaliações	8
Visualização da Informação	8

Conteúdo Programático - EAD

Tópico	Carga horária
Acompanhamento de preparação para seminário	6
Acompanhamento de escrita de artigo	6

Total

60

Metodologia de Ensino

Presencial

- Aulas expositivas/discursivas
- Leituras e estudos complementares (artigos, trabalhos, documentações, ferramentas, etc.)
- Participação em aula: pesquisas propostas, discussões/apresentações de trabalhos/artigos aplicados, exercícios
- Seminários

EAD

- Planejamento de seminário de pesquisa em tema da disciplina usando a plataforma Google classroom. Acompanhamento para cada discente e entrega de milestones em datas pré estabelecidas.
- Planejamento e acompanhamento de escrita de artigo a respeito de tema definido, aliado ao seminário utilizando a ferramenta do Google classroom. Entregas de etapas do artigo segundo deadlines pré-definidos por meio da plataforma.

Recursos Didáticos

Quadro, projetor, ferramentas de software open source, periódicos, artigos, acesso à Internet.

Critérios de Avaliação

A avaliação dar-se-á com base em três marcos: (i) Pesquisas e trabalhos/exercícios passados ao longo da disciplina (**T1**); (ii) Seminário sobre tópico de pesquisa da disciplina e tema de mestrado do discente, com apresentação oral e escrita (**T2**); (iii) Artigo escrito e submetido resultante da pesquisa realizada vinculada à disciplina e ao tema de mestrado do discente (**T3**). Todas as atividades terão sub marcos e deadlines definidos a serem cumpridos. Ao final, a nota da disciplina será **Nota = (T1 + T2 + T3)/3**.



Bibliografia

1. ELMASRI, R.; NAVATHE S. B. Sistemas de Banco de Dados. 6 ed. São Paulo: Pearson Education, 2011.
2. HAN, J.; KAMBER, M.; PEI, J. Data Mining Concepts and Techniques. 3. ed. Waltham (USA): Morgan Kaufmann, 2012.
3. KIMBALL, R.; ROSS, M. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. 2. ed. [S.l.]: Wiley, 2002.
4. SHAHBAZ, Q. Data Mapping for Data Warehouse Design. 1. ed. [S.l.]: Morgan Kaufmann, 2015.
5. DONG, X. L.; SRIVASTAVA, D. Big Data Integration. [S.l.]: Morgan & Claypool. 2015.
6. TAN P., STEINBACK M. e KUMAR V. Introduction to Data Mining, Pearson, 2013.
7. GOLDSCHMIDT R., PASSOS E., Datamining - Conceitos, Técnicas, Algoritmos, Orientações e Aplicações. Editora Campus, 2015.
8. AMARAL, F. Introdução à Ciência de Dados. Editora Alta Books, 2016.
9. Artigos relacionados aos tópicos de interesse na disciplina, selecionados a partir de artigos publicados em revistas e eventos de alto impacto.

Observações

(Nenhuma)