

FICHA 2 - PLANO DE ENSINO

CÓDIGO: SIN195	DISCIPLINA: MINERAÇÃO DE DADOS		TURMA: A			
NATUREZA: Obrigatória			MODALIDADE: Parcial EAD - 20% EAD			
CH TOTAL: 60h			CH Prática como Componente Curricular (PCC): 0h		CH Atividade Curricular de Extensão (ACE): 0h	
Padrão (PD): 30h	Laboratório (LB): 30h	Campo (CP): 0h	Orientada (OR): 0h	Estágio (ES): 0h	Prática Específica (PE): 0h	Estágio de Formação Pedagógica (EFP): 0h
FICHA 2 PREENCHIDA PELO DOCENTE: DENISE FUKUMI TSUNODA						

Criação: 25/6/2023

Modificação: 25/6/2023

EMENTA

Aspectos teóricos e práticos da descoberta de conhecimento em bancos de dados utilizando tarefas e técnicas de mineração de dados.

JUSTIFICATIVA PARA OFERTA EAD:

A oferta da disciplina na modalidade a distância foi aprovada conforme ajuste curricular e atende à distribuição de 20% da carga horária do curso prevista no Projeto Pedagógico, com base na LDB e na Resolução nº 72/10 CEPE, com o objetivo de estimular as e os discentes no processo de aprendizagem. A oferta de carga horária a distância também está pautada no artigo 2º do Decreto nº 9.057 de 25 de maio de 2017, e na Lei nº 13.146/2015 - Lei Brasileira de Inclusão da Pessoa com Deficiência. Assim, está alinhada com a promoção da acessibilidade e inclusão, conforme orientado pela Superintendência de Inclusão, Políticas Afirmativas e Diversidade (SIPAD) da UFPR. Somam-se a isso, ainda, as estratégias para assegurar a acessibilidade digital tecnológica e educacional disponíveis na UFPR VIRTUAL, o ambiente virtual de aprendizagem (AVA) da UFPR.

PROGRAMA

Tópico 1 – Conceitos fundamentais: Inteligência Artificial, Computação Natural, Análise de Dados e Mineração de Dados

Tópico 2 – Tarefas de mineração de dados: classificação, associação e agrupamento

Tópico 3 – Pré-processamento de base de dados: discretização, limpeza, transformação e outros.

Tópico 4 – Análise de resultados: matriz de confusão, taxa de acerto, confiança, suporte, distâncias e outros



Tópico 5 – Aplicação prática dos tópicos anteriores

Semana	Data	Carga horária			Conteúdo	Metodologia	Ambiente & Ferramentas	ENTREGAS
Presencial	EaD	Total	1	25/07/2023 27/07/2023	2	2	4	Apresentação da disciplina, equipes, trabalhos, ferramentas
Aula expositiva e definição da AT1	Sala de aula, UFPR Virtual, TEAMS, DSA, Weka	01/ago	2	01/08/2023 03/08/2023	4	0	4	Conceitos básicos
Apresentações - IGNITE Dúvidas pelo Teams	Sala de aula, UFPR Virtual, Office 364		3	08/08/2023 10/08/2023	4	0	4	KDD e Mineração de Dados; Tarefa de Classificação: ID3 recursivo
Apresentações IGNITE Aula expositiva: ID3: AT2	Sala de aula, UFPR Virtual, Office 365	22/ago	4	15/08/2023 17/08/2023	4	0	4	Tarefa de classificação: C4.5
Aula expositiva: AT5	Sala de aula, UFPR Virtual, Office 365	29/ago	5	22/08/2023 24/08/2023	2	2	4	Inteligência Artificial e Ciência de Dados



DSA: AT3	Office 365, UFPR Virtual, DSA, Weka	05/set	6	29/08/2023 31/08/2023	2	2	4	Tarefa de Classificação WEKA e Excel
DSA: AT4 AT6	Laboratório, Sala de aula, UFPR Virtual, Office 365, Weka, DSA	12/set	7	05/09/2023	2	0	2	Primeira Avaliação Individual
	Sala de aula		8	12/09/2023 14/09/2023	4	0	4	Trabalho Final Segunda Avaliação Individual
TF	Laboratório, Office 365, UFPR Virtual, Weka Sala de aula		9	19/09/2023 21/09/2023	4	0	4	Tarefa de Associação: Apriori
Aula expositiva: Apriori: AT7	Sala de aula, UFPR Virtual, Office 365	03/out	10	26/09/2023 28/09/2023	2	2	4	Tarefa de Associação WEKA e Excel
AT8	Office 365, UFPR Virtual, Weka	10/out	11	03/10/2023 05/10/2023	4	0	4	Tarefa de agrupamento: K-Means
Aula expositiva: AT9	Sala de aula, UFPR Virtual, Office 365	24/out	12	10/10/2023	2	0	2	Terceira Avaliação Individual



	Sala de aula		13	17/10/2023 3 19/10/2023 3	2	2	4	Trabalho Final
TF	Laboratório, Office 365, UFPR Virtual, Weka	05/nov	14	24/10/2023 3 26/10/2023 3	2	2	4	Trabalho Final
TF	Laboratório, Office 365, UFPR Virtual, Weka		15	31/10/2023 3	2	0	2	Trabalho Final
TF	Laboratório, Office 365, UFPR Virtual, Weka		16	07/11/2023 3 09/11/2023 3	4	0	4	Apresentações dos trabalhos
Aula expositiva	Sala de aula, Office 365, UFPR Virtual, Weka		17	14/11/2023 3	2	0	2	Fechamento
	Sala de aula				48	12	60	
Exame final da disciplina: 05/12/2023								

OBJETIVO GERAL

O aluno deverá conhecer e ser capaz de aplicar diferentes técnicas estatísticas e computacionais de exploração de dados na identificação de padrões relevantes e não óbvios em bancos de dados.

OBJETIVOS ESPECÍFICOS



O aluno deverá ser capaz de:

- 1 discutir os conceitos básicos relacionados à mineração de dados;
- 2 aplicar e explicar resultados das principais tarefas de mineração de dados;
- 3 justificar a relevância desta disciplina;
- 4 identificar e compreender aplicações práticas da mineração de dados.

PROCEDIMENTOS DIDÁTICOS

Aulas, teórico práticas, expositivo dialogadas em encontros PRESENCIAIS nos quais a docente faz exposição do tema do programa, discute com os alunos os métodos e aplicações e propõe exercícios e atividades em grupo ou individuais, além das atividades avaliativas.

Serão utilizados como recursos: sala de aula no SA, laboratório de computadores, sala no ambiente virtual de aprendizagem – AVA, mais especificamente o UFPR Virtual; o Microsoft Office 365 disponível institucionalmente, o site da Data Science Academy como suporte ao conteúdo e o software Weka. Para as eventuais aulas não presenciais o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino à distância.

Como suporte ao ensino será utilizada a infraestrutura da UFPR tais como: software (Office 365) e AVA.

Em atendimento à Resolução 72/10-CEPE, serão observados os seguintes aspectos:

- a) sistema de comunicação – encontros presenciais, AVA (UFPR Virtual), correios eletrônicos trocados entre coordenação e alunos e professor e alunos;
- b) modelo de tutoria a distância e presencial – não está prevista a tutoria nesta disciplina. As dúvidas serão esclarecidas pela docente e, quando existir, pelos alunos do programa de Monitoria ou alunos de PPGs em estágio docência;
- c) material didático específico – o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino presencial;
- d) infraestrutura de suporte tecnológico, científico e instrumental à disciplina – o curso conta com Ambiente Virtual de Aprendizagem específico e pacote Office 365, incluindo sala no Teams para eventuais esclarecimentos e interação;
- e) previsão de período de ambientação dos recursos tecnológicos a serem utilizados pelos discentes: os alunos recebem capacitação no Ambiente de Aprendizagem Virtual no primeiro semestre do Curso e esta é uma disciplina optativa, prevista para oferta nos últimos períodos do curso. Assume-se que os alunos já tenham domínio no uso das tecnologias.

TODAS AS AVALIAÇÕES (PROVAS) SERÃO INDIVIDUAIS E PRESENCIAIS.

As atividades podem ser presenciais ou remotas, conforme definição em sala de aula e rubricas.

FORMAS DE AVALIACAO



Algumas etapas avaliativas serão assíncronas e outras presenciais. Os critérios das atividades e avaliações estão definidos e serão publicados no ambiente virtual, no primeiro dia de aula, para todos os alunos. As três notas principais para composição da média final (cálculo da média aritmética simples das três notas), são:

1 primeira nota: média aritmética das atividades (AT) realizadas individualmente ou em equipe (conforme solicitado) com avaliação conforme rubricas publicadas no início da disciplina;

2 segunda nota: média aritmética das duas avaliações individuais (AV) com consulta;

3 terceira nota: trabalho (TF) **em dupla** com **avaliação conforme rubrica "TRABALHO FINAL"** divulgada no primeiro dia de aula que considerará:

a) qualidade do levantamento bibliográfico para o referencial teórico;

b) explicação do funcionamento do algoritmo/heurística recebido;

c) aplicações práticas (exemplos de uso em outros contextos) do algoritmo recebido;

d) profundidade na abordagem do método (implementado ou em pacote);

e) adaptação da base de dados recebida à implementação realizada ou pacote utilizado;

f) relatório detalhado do experimento;

g) análises dos resultados obtidos;

h) comparação dos resultados com a outra equipe com o mesmo método;

i) registros (atas) das reuniões;

j) criatividade da apresentação das equipes;

k) uso de normas para elaboração de trabalhos de pesquisa (normas ABNT);

l) desempenho individual e coletivo dos membros das equipes durante a apresentação EM BANCA constituída por dois ou três professores.

As frequências no AVA serão calculadas conforme a entrega das atividades e avaliações propostas.

As demais frequências serão controladas por chamada em sala de aula.

O limite de ausências permitido é 25% da carga horária da disciplina.

BIBLIOGRAFIA BÁSICA

CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. **Introdução à mineração de dados**: conceitos básicos, algoritmos e aplicações. São Paulo: Editora Saraiva, 2016. 978-85-472-0100-5. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/>. Acesso em: 22 jun. 2023.

NG, Andrew. **Machine learning learning**. Disponível em: <https://www.deeplearning.ai/machine-learning-yearning/>. Acesso em: 22 jun. 2023.

SILVA, Leandro Augusto da; PERES, Sarajane Marques; BOSCARIOLI, Clodis. **Introdução à mineração de dados** - com aplicações em R. Rio de Janeiro: Grupo GEN, 2016. 9788595155473. Disponível em:



<https://integrada.minhabiblioteca.com.br/#/books/9788595155473/>. Acesso em: 22 jun. 2023.

Vídeos

RAMESH, Raj. **What is artificial intelligence?** In 5 minutes. Disponível em: <https://www.youtube.com/watch?v=2ePf9rue1Ao&t=155s>. Acesso em: 22 jun. 2023.

FERREIRA, Eduardo Vargas. **Aula 6 – Métodos de reamostragem**. Disponível em: <https://www.youtube.com/watch?v=pqeYQahl1Rk&list=PLWPcjBWc0bg7MUc0tFquAMQX1kA971ubc&index=6>. Acesso em: 22 jun. 2023.

BIBLIOGRAFIA COMPLEMENTAR

URKOV, Andriy. **The hundred-page machine learning book**. Disponível em: <http://themlbook.com/wiki/doku.php>. Acesso em: 22 jun. 2023.

DEISENROTH, Marc Peter; FAISAL, A. Aldo e ONG, Cheng Soon. **Mathematics for machine learning**. Disponível em: <https://mml-book.github.io>. Acesso em: 22 jun. 2023.

GOOGLE. **Machine learning crash course**. Disponível em: <https://developers.google.com/machine-learning/crash-course/>. Acesso em: 22 jun. 2023.

POOLE, David L.; MACKWORTH, Alan K. **Artificial intelligence: foundations of computational agentes**. 2ª ed. Disponível em: <https://artint.info/2e/html/ArtInt2e.html>. Acesso em: 22 jun. 2023.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding machine learning**. Disponível em: <https://www.cse.huji.ac.il/~shais/UnderstandingMachineLearning/>. Acesso em: 22 jun. 2023.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement learning: an introduction**. 2ª ed. Disponível em: <http://incompleteideas.net/book/RLbook2018.pdf>. Acesso em: 22 jun. 2023.

CRONOGRAMA DE AULAS

Não disponível

